

Image-based Viewpoint Classification Related to the Bus-Waiting for Assisting the Blind

著者	Tangsuksant Watcharin
その他のタイトル	視覚障害者のバス利用支援のための画像処理を用いたカメラ視点識別に関する研究
学位授与年度	令和元年度
学位授与番号	17104甲生工第356号
URL	http://hdl.handle.net/10228/00007482



Image-based Viewpoint Classification Related to the Bus-Waiting for Assisting the Blind

視覚障害者のバス利用支援のための画像処理を用いたカメラ視点識別に関する研究

WATCHARIN TANGSUKSANT

ワッチャリン タンスサント

Kyushu Institute of Technology
Graduate School of Life Science and Systems Engineering

September, 2019

Image-based Viewpoint Classification Related to the Bus-Waiting for Assisting the Blind

by
Watcharin Tangsuksant
Student ID number: 16899026

Supervisor: Prof. Chikamune Wada

Abstract

The bus identification using smartphone camera is one useful application for blind people who travel independently in daily life. Although, many existing researches have focused on the bus identification by using image processing, those researches did not concern the viewpoint of image before the oncoming bus appear in the image. This research proposes the definition and classification of suitable viewpoint of bus-waiting for aiding blind people, which is proposed into three conditions as following: 1) non-congested traffic; 2) congested traffic; and 3) obstacles detection along the road. The first condition of non-congested traffic classified the viewpoint using the road area consideration, which this research applied the Rotational Invariant of Uniform of Local Binary Pattern technique for extracting the road area, and the Back-propagation of Artificial Neural Network for the viewpoint classification. The second condition is congested traffic, which appears a huge number of car in the image. The distribution of car in the image was calculated by many features, which the optimized results showed that seventeen selected features and Random Forest classifier provided the high performance. For third condition, the obstacles along the road will be consideration in case of non-congested traffic. This research combined many existing techniques of image processing to detect the obstacles along the road, which consisted of two main processes: 1) obstacle's location detection and 2) obstacle's height estimation. The proposed detected technique can implement in the daylight condition with high performance. According to experimental results, the high performance have shown by 98.56%, 86.00% for non-congested and congested traffics, respectively. Moreover, the performance of obstacles position detection and height estimation were shown by 91.20%, 86.00%, respectively. Based on these results, these are feasible to apply for viewpoint classification in order to assist blind people, who are independently waiting for the bus.

This page intentionally left blank

Acknowledgement

First and foremost, I would like to thank to the Japanese Government for full scholarship (MEXT) during my doctoral course. Moreover, the committee of Global Advanced Assistive Robotics Course (GAAR), Kyushu Institute of Technology, who selected me to get this scholarship. It is one of the best opportunity in my life to receive this scholarship.

I would like to express my appreciation to my supervisor Professor Chikamune WADA who are coaching and mentoring of my research. Without his guidance and impetus, my dissertation would not have been successful. Moreover, I would like to express my appreciation to supervisory committee who give me suggestions to improve the dissertation: Professor Kiyohisa NATSUME, Professor Hiroyuki MIYAMOTO, Professor Takeshi SAITOH and Professor Chikamune WADA.

To my laboratory seniors and colleague members: Dr. Romy Budhi Widodo, Dr. Takayuki Nagasaki, Dr. Jin Fang, and Dr. Ahmed Almassri who are kind seniors, and likewise to be my brother and sister. My colleagues: Mr. Yushiki Nishisako and Mr. Tsuyoshi Uezono who picked me up at the Fukuoka airport on first date that I arrived in Japan. I many thanks for everything that you help me. In addition, I really thank to Mr. Kodai Kitagawa, Mr. Kiskeya Iwasaki and Mr. Masashi Noda who always discuss and help me about academic research and some problems for my daily life. For all colleague members: Mr. Takeshi Jouyashiki, Mr. Hiroyuki Kuraoka, Mr. Ibai Gorordo Femandez, Mr. Shogo Okamatsu, Mr. Shunsuke Fukumoto, Mr. Yejong Lee, Mr. Yunchu Huang, Mr. Hideyuki Tobata, Mr. Hiro Ohnishi, Ms. Akari Tokuda, Mr. Ryoya Nakamura, Mr. Koji Matsumoto, and Mr. Koji Sakamoto, all the best for you all and success for your study and career in the next future.

I would like to appreciate to Associate Professor Theekapun Charoenpong who inspired me to study the doctoral course. Moreover, I appreciate to Associate Professor Chuchart Pintavirooj and Dr. Theerasak Chanwimalueang, who were advisors of my master and bachelor courses, for knowledge and research skill that you gave me. In addition, I appreciate to all teachers that taught me for academic skill.

Finally, many thanks to my beloved family: Mr. Wichean Tangsuksant (Dad), Mrs. Anchalee Tangsuksant (Mom), Mr. Akekachai Tangsuksant (Brother), Mr. Thawatchai Tangsuksant (Brother), Mrs. Rarinthon Tangsuksant (Sister-in-law), Ms. Evarin Tangsuksant (Niece), Mr. Archawin Tangsuksant, and Ms. Darikarn Paramee (Girlfriend) with her family, they are always my encouragement and support me along the way.

Watcharin Tangsuksant, September 2019

Contents

Abstract.....	i
Contents	v
List of Figures.....	viii
List of Tables	xii
1. Introduction	1
1.1 Background of Study	1
1.2 Purposes of the Study.....	3
1.3 The Present Configuration of the Paper.....	3
2. Previous Studies	8
2.1 Related Previous Research.....	8
2.1.1 Public Transportation and Assistive System for Blind People	8
2.1.2 Bus Identification System for Blind People	9
2.2 Problem of Existing Previous Researches for Bus Identification System	11
2.3 The Proposed Method	13
2.3.1 Proposed Method for Non-Congested Traffic	13
2.3.2 Proposed Method for Congested Traffic	13
2.3.3 Proposed Method for Obstacle Detection along the Road	14
3. Suitable Viewpoint Definition of Waiting for the Bus	16
3.1 The Possible Image Viewpoints of Waiting for the Bus Using Smartphone	16
3.1.1 Excessive High-Angle of Viewpoint	17
3.1.2 Excessive Low-Angle of Viewpoint.....	17
3.1.3 Excessive Right Direction of Viewpoint	18
3.1.4 Excessive Left of Viewpoint.....	18
3.1.5 Suitable Angle and Direction of Viewpoint.....	18
3.2 A Suitable Viewpoint Definition of Waiting for the Bus using Smartphone	19
3.2.1 Suitable Tilt Definition of Smartphone Use while Waiting for the Bus	19
3.2.2 Suitable Panning Definition of Smartphone Use while Waiting for the Bus ..	21
3.3 Summary	30
4. Classification of Viewpoints while Waiting for the Bus in Situation of Non-Congested Traffic	32

4.1	Objective of Classification of Viewpoints in case of Non-Congested Traffic	32
4.2	Proposed Method for Classification of Viewpoint based on Non-Congested Traffic	33
4.2.1	Finding the Reference of Road Area	34
4.2.2	Pre-Viewpoint Classification	39
4.2.3	Road Segmentation and Feature Extraction	41
4.2.4	Viewpoints Classification based on Non-Congested Traffic	48
4.3	Experiments and Results	48
4.3.1	Performance Testing for Selecting the k_n by using ANN Classification	49
4.3.2	Accuracy of Pre-Viewpoint Classification using Criteria Setting	50
4.3.3	Feature Extraction Performance	51
4.3.4	Viewpoint Classification Accuracy	52
4.3.5	Comparison Results	54
4.4	Discussion	55
4.5	Summary	59
5.	Classification of Viewpoints while Waiting for the Bus in Situation of Congested Traffic	61
5.1	Objective of Classification of Viewpoints in case of Congested Traffic	61
5.2	Viewpoint Estimation in case of Congested Traffic based on Suitable Viewpoint Definition	62
5.3	Proposed Method for Classification of Viewpoint based on Congested Traffic	63
5.3.1	Car Detection by YOLO	64
5.3.2	Data Normalization	65
5.3.3	Car Distribution-Feature Extraction	67
5.3.4	Viewpoints Classification based on Congested Traffic	69
5.4	Experiments and Results	71
5.4.1	Car-detection Performance	71
5.4.2	Optimization for Features and Classifiers Selection	72
5.4.3	Performance Measurement for Viewpoint Classification	74
5.4.4	Comparison Result	75
5.5	Discussion	75
5.6	Summary	76
6.	Obstacle Detection along the Road	78

6.1	Objective of the Obstacle Detection along the Road	78
6.2	Properties of Obstacles	79
6.2.1	Range of Interest.....	79
6.2.2	Height of Obstacle	79
6.3	Proposed Method for the Obstacle Detection	81
6.3.1	The Detection of Obstacle's Position along the Road.....	81
6.3.2	Estimation of Height of Obstacles.....	89
6.4	Experiments and Results.....	92
6.4.1	Detection Performance of Obstacle's Position along the Road.....	92
6.4.2	Optimization of Color Moment Technique for Obstacle's Height Estimation	93
6.5	Discussion	99
6.5.1	Discussion of Obstacle's Position Detection.....	99
6.5.2	Discussion of Optimization Process for Obstacle's Height Estimation	99
6.6	Summary	102
7	Conclusions and Future Work	104
7.1	Conclusion.....	104
7.2	Future Work	105
	References	108
	Appendix Published Papers	114

List of Figures

Figure 1.1 General bus number recognition for existing systems using image-processing technique.....	2
Figure 1.2 Dissertation chapter overview.	5
Figure 1.3 Overview of proposed method related to each chapter.	6
Figure 2.1 Example of two main bus identification system for blind people.....	9
Figure 2.2 Example of image viewpoints for the real situation.....	12
Figure 3.1 Five main possible image viewpoints of waiting for the bus.....	16
Figure 3.2 Example of other image viewpoints.....	17
Figure 3.3 Two main factors of suitable viewpoint appearance while waiting for the bus. .	19
Figure 3.4 Holding smartphone and its tilt.	20
Figure 3.5 Suitable of tilt-down ($y = 9.5 \text{ m/s}^2$ and $z = 3.5 \text{ m/s}^2$ of accelerometer value). ...	20
Figure 3.6 Suitable of tilt-up ($y = 9.0 \text{ m/s}^2$ and $z = -4.0 \text{ m/s}^2$ of accelerometer value).....	21
Figure 3.7 Tilt-down and tilt-up of smartphone with x, y and z of accelerometer values....	21
Figure 3.8 Camera panning.	22
Figure 3.9 Different types of Thai-public bus and bus number size.	23
Figure 3.10 Simulated route bus number in Thailand by 15 x 15 board size.	23
Figure 3.11 Digital LUX meter for our study.	25
Figure 3.12 Example of different capturing distance and its image output.....	26
Figure 3.13 Processing step of OCR testing by capturing the simulated bus number.....	27
Figure 3.14 Estimation of viewpoints with/without oncoming bus.....	29
Figure 3.15 A suitable viewpoint with 25% of parameter setting.	29
Figure 4.1 Different patterns of road area with different illuminations between daytime and nighttime.....	32
Figure 4.2 Block diagram of the proposed method.	33
Figure 4.3 $LBP_{P,R}^{riu2}$ with $P = 8$ and $R = 1$	35
Figure 4.4 Example of transformed image by $LBP_{8,1}^{riu2}$	35
Figure 4.5 Step of histograms extraction for each sliding sub-window.	36

Figure 4.6 Five different $LBP_{8,1}^{riu2}$ histogram representations from k -means.	37
Figure 4.7 Basic structure of Artificial Neural Network.	38
Figure 4.8 The suitable viewpoint with excessive left panning.	39
Figure 4.9 Comparison of excessive left panning viewpoint and other viewpoint patterns.	40
Figure 4.10 Comparison of plotting features between large road area viewpoints and other viewpoints.	41
Figure 4.11 The step of road edge detection with Hough Line Transformation method.	42
Figure 4.12 Line touching on the circle's edge.	43
Figure 4.13 Upper and Lower centroid separation for two-line detection.	44
Figure 4.14 Case of one-line detection.	44
Figure 4.15 Setting cut-off value for noise elimination.	45
Figure 4.16 Example of gray-scale histogram with different illumination on the reference road area.	46
Figure 4.17 Gray-Scale and Seed Region Growing Segmentation.	46
Figure 4.18 The necessary road-features for viewpoints classification.	47
Figure 4.19 Eight different conditions between daytime and nighttime for experiment.	49
Figure 4.20 Graphical representation of error percentage for feature extraction.	56
Figure 4.21 Compared graphical between viewpoint classification and error of feature extraction.	56
Figure 4.22 Comparison graph of three different methods of our proposed.	57
Figure 4.23 Example of error case for line detection.	58
Figure 4.24 Many detected lines of HLT process.	59
Figure 5.1 Difference between non- and congested traffic viewpoint.	61
Figure 5.2 Estimation of suitable and unsuitable viewpoint in case of congested traffic based on suitable definition.	62
Figure 5.3 Congested traffic viewpoints under daytime and nighttime.	63
Figure 5.4 Four main processes of proposed method for viewpoint classification in case of congested traffic.	63
Figure 5.5 The architecture of convolutional layers neural network for YOLO.	64

Figure 5.6 Outcomes of car detection using the YOLOv2 technique (these images are related to the graphs of Figure 5.7).	65
Figure 5.7 Data normalization.	66
Figure 5.8 Feature-matrix arrangement and its labels.	71
Figure 6.1 Obstacles along the road.	78
Figure 6.2 Range of Interest for obstacles consideration.	79
Figure 6.3 Horizon line and vanishing point of perspective image.	80
Figure 6.4 Horizon line and oncoming bus.	80
Figure 6.5 Block diagram of the detection of obstacle's position.	81
Figure 6.6 The process of image transformation using Rotational invariant of uniform local binary patterns.	82
Figure 6.7 Image transformation from k -means clustering.	83
Figure 6.8 Five different histogram representation of $LBP_{8,1}^{riu2}$ using k -means.	84
Figure 6.9 Road line detection and initial road area selection (first criterion).	85
Figure 6.10 Road line detection and initial road area selection (second criterion).	85
Figure 6.11 The line selection using line touching the circle.	86
Figure 6.12 Detection of road line and road segmentation.	87
Figure 6.13 Vertical projection of each line scan (S_n) on binary image.	87
Figure 6.14 Normalized signal of vertical projection.	88
Figure 6.15 Example of obstacle detection using the proposed method.	89
Figure 6.16 Concept for estimating the height of obstacle.	89
Figure 6.17 The concept of the proposed method and an ideal graph for data clustering.	91
Figure 6.18 Example of matching and non-matching images.	94
Figure 6.19 Comparison graph for data clustering between HSV and RGB color model using the weight of 1, 1, 2.	95
Figure 6.20 Comparison graph of transpose matrix (H' , I') between the weight of 3, 1, 1 and 3, 1, 2.	98
Figure 6.21 Accuracy comparison between non-transpose (H , I) and transpose (H' , I') matrices.	100

Figure 6.22 Example of obstacles detection using our proposed method.	101
Figure 7.1 Overview of proposed system and future works.....	106
Figure 7.2 Example image of straight and curved roads.	107

List of Tables

Table 3.1 Characters appearing on the Thai-route bus number.	24
Table 3.2 Specification of digital LUX meter (LX-1332B).	25
Table 3.3 Possibility of route bus number recognition using OCR technique.	27
Table 4.1 List of features extracted with their significance.	48
Table 4.2 Performance testing for selecting the group of k -means by using ANN classification.	50
Table 4.3 Accuracy of pre-viewpoint classification using criteria setting.	50
Table 4.4 Performance testing for feature extraction.	52
Table 4.5 Accuracy of viewpoint classification.	53
Table 4.6 Comparison of three different combined methods based on our proposed algorithm.	54
Table 4.7 Comparison result between our proposed method and related research.	55
Table 5.1 Parameter setting of YOLOv2 for car detection.	65
Table 5.2 List of features and their descriptions.	67
Table 5.3 Initial parameter setting for simple-decision tress.	69
Table 5.4 Initial parameter setting for Random forest.	69
Table 5.5 Initial parameter setting for Naïve Bayes	70
Table 5.6 Initial parameter setting for multi-layer perceptron.	70
Table 5.7 Initial parameter setting for Support-vector machine	70
Table 5.8 Car-detection performance using YOLO and the proposed parameter setting.	72
Table 5.9 Comparison of accuracy with different feature selections and classifiers.	73
Table 5.10 Confusion matrix for viewpoint classification.	74
Table 5.11 Comparison result between our proposed method (congested traffic) and related research.	75
Table 6.1 Confusion Matrix of Obstacle Detection using The Proposed Method.	92
Table 6.2 The $-SD'$ and $+SD$ Comparison between HSV and RGB.	95
Table 6.3 Performance of RGB Color Model Testing based on Non-Transpose Matrix (H, I).	97

Table 6.4 Performance of RGB Color Model Testing based on Transpose Matrix (H' , I').	97
Table 6.5 Comparison of Distance Error Measurements between Weight of (3, 1, 1) and (3, 1, 2).	99

This page intentionally left blank

1. Introduction

1.1 Background of Study

According the World Health Organization, the number of completely blind people, no light perception (NLP), was 36 million in 2017 [1-2]. Blind people encounter several problems in their daily life, such as reading texts or warning signs in public places. While these problems might be addressed in developed countries, many underdeveloped and developing countries still have no available systems to support the blind population.

There have been several tools and devices assisting the blinds. The basic necessary devices are white cane, taking clock and portable braille typing. Moreover, the advance of technology, nowadays, is improving a huge number of useful devices for helping them. For example, WeWALK [3] is a smart cane for visually impaired to see the obstacles in front of them. In addition, the intelligent glasses have been developed by many organizations and famous companies such as OrCam MYEye [4], OpenGlass [5] and SIGHT [6]. These glasses can let the blind know what is going on in front of them, detect and recognize the many objects. Smartphones, especially, become an important device for everyone including blind people. There have been many assistive applications containing on the smartphones for assisting the blind people. For example, VoiceOver [7] on iPhone and iPad is becoming a standard application for helping blind to use their smartphones. This application can read the context on a display to navigate and select their options.

Especially, the bus identification system is still a problem that the blind people have to face on their travels. For example, the public bus transportation, in Thailand, has no systems to support disabled people waiting for buses, and the estimation of bus arrival times is impossible for everyone. There are no announcements regarding the details of bus routes and oncoming buses either; therefore, users have to observe the bus number by themselves at the bus stop, and then wave their hand to signal the bus driver to pick them up. Although, the blind people can reach the bus stop by themselves but it is still difficult to get the oncoming bus. According to the interviews of some Thai blind people [15], at the King Chulalongkorn Memorial Hospital's Low Vision Clinic, many of them need care-takers to help them for their

activities, some young aged of them can take the smartphone to do something independently such as reading sign. However, the travelling independently by public buses is still one of major difficulty for them. Although, blind individuals who travel independently can obtain help from neighboring people at the bus stop, blind individuals, unfortunately, often face extreme difficulty when they stay alone or have no one to assist them at the bus stop.

The bus identification is an application for aiding blind individual to know the oncoming bus. Although, many systems of bus identification have been proposed by previous studies, those systems could be separated into two main types. The first type uses the transceiver communication [8-14] to identify the oncoming bus such as Bluetooth, RFID, WiFi and Global Positioning System (GPS). The second type [15-19] uses the image processing to capture and recognize the bus number for an oncoming bus using the static camera at bus stop or user's smartphone. Although a transceiver communication system is a good idea to identify the bus for blind individuals, this communication system is quite large for practical use because transceiver modules have to be installed for every bus station, buses, and user. Further, system maintenance is difficult when some part of the module is not working. Therefore, we attend to the image processing system for bus identification using smartphone, because it is portable and comfortable for blind people recently.

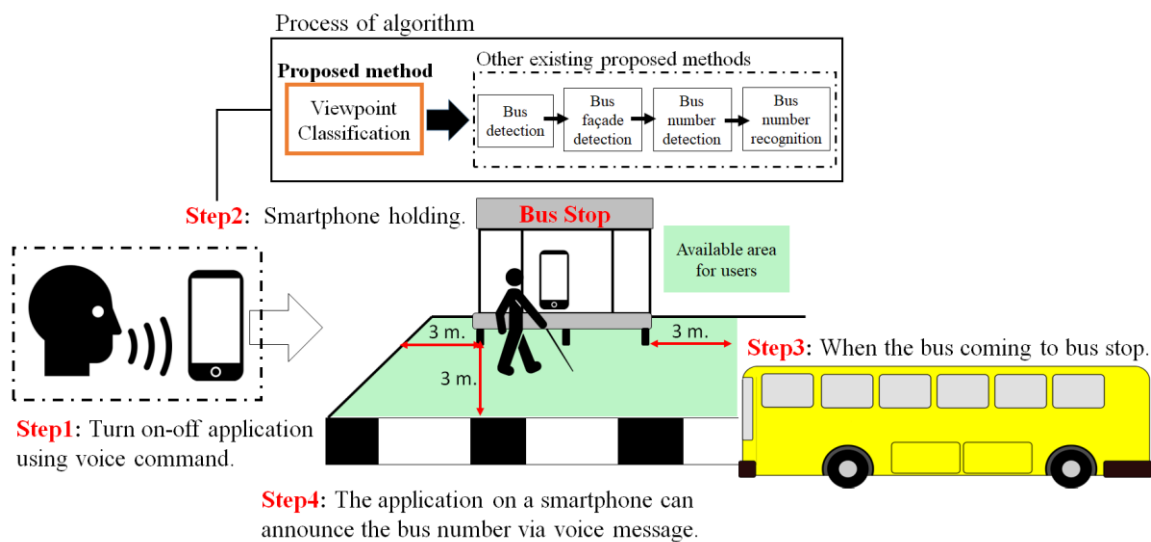


Figure 1.1 General bus number recognition for existing systems using image-processing technique.

Generally, the bus identification using smartphone to capture and recognize the oncoming bus number consist of four steps for existing proposed algorithm. First, the system will detect the position of oncoming bus in the image, and then finding the bus façade for region of interest. After that, the bus number is respectively detected and recognized. According to the existing proposed processes, many researchers have already presented the various techniques. However, there are no previous researches concerning the suitable viewpoints of bus-waiting. Practically, when the blind people uses their smartphone to capture the images, they cannot know the captured viewpoints by themselves. Figure 1.1 shows the overview of our own system design that consists of four steps. Firstly, a voice command is used for turning the application on or off. Subsequently, users use their smartphone to take a video when waiting for an oncoming bus at the bus stop, which the available area is defined by 3 meters next to road area, and around bus stop. Finally, the application on smartphone will convey the bus number via a voice announcement to the user. Although, the bus stops in the city have two different types such as the typical bus stop pole and bus stop with bench and roof, the users should wait the bus around the available area until the bus comes.

1.2 Purposes of the Study

Based on the above background, the objective of this study is to propose the suitable viewpoint definition and classification of bus-waiting for assisting blind people. Since, there are various conditions and factors of the real scenarios, three situations for bus-waiting are considered by this study. The list of details of our proposed situations as follows:

1. The situation of non-congested traffic in daytime and nighttime conditions.
2. The situation of congested traffic in daytime and nighttime conditions.
3. The situation of obstacles appearance along the road.

1.3 The Present Configuration of the Paper

This paper is organized in seven chapters, which shows in Figure 1.2. The content of each chapter is as follows:

Chapter 1 explains the study's background with a brief literature review and discusses

the purposes of the study.

Chapter 2 presents the related previous studies, discusses the problems, and proposes the methods for solving those problems.

Chapter 3 explains the suitable viewpoints definition for bus-waiting.

Chapter 4 proposes an idea to classify the viewpoints of bus-waiting in case of non-congested traffic without obstacle along the road. This chapter is divided into five parts: 1) objective of classification of viewpoint in case of non-congested traffic, 2) proposed method for classification of viewpoint, 3) experiments and results, 4) discussion, and 5) summary.

Chapter 5 proposes an idea to classify the viewpoints of bus-waiting in case of congested traffic without obstacle along the road. This chapter is divided into six parts: 1) objective of classification of viewpoint in case of congested traffic, 2) viewpoint estimation in case of congest traffic related to suitable viewpoint definition, 3) proposed method for classification using car distribution features, 4) experiments and results, 5) discussion, and 6) summary.

Chapter 6 proposes an idea to detect the obstacles along the road. This chapter is divided into six parts: 1) objective of the obstacle detection, 2) properties of obstacles, 3) proposed method for the obstacle detection, 4) experiments and results, 5) discussion, and 6) summary.

Chapter 7 includes the conclusion in addition to future work direction.

The outline of this dissertation is as follows:

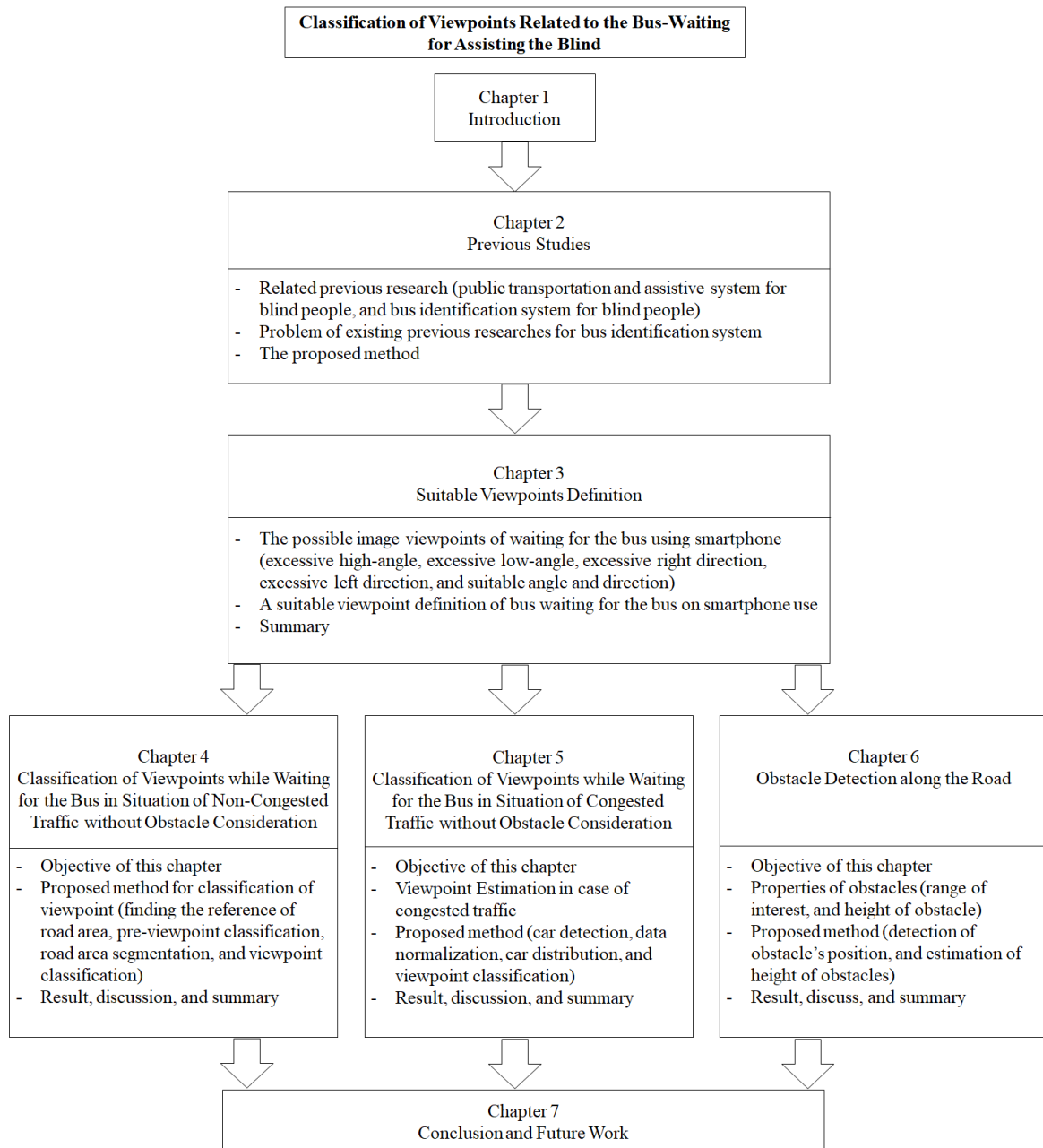


Figure 1.2 Dissertation chapter overview.

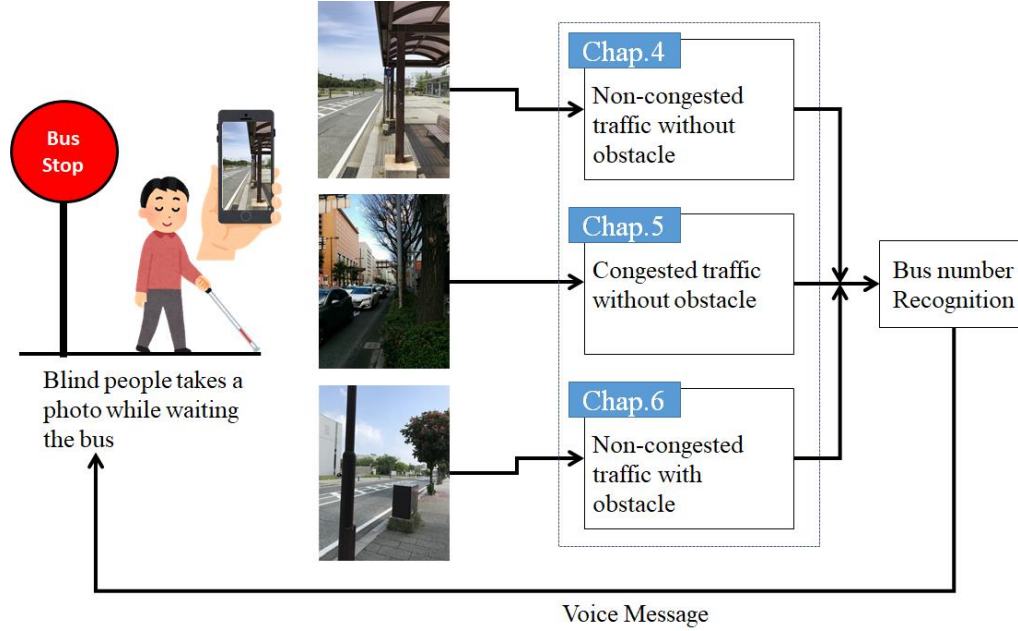


Figure 1.3 Overview of proposed method related to each chapter.

To achieve the objective for viewpoint classification, Figure 1.3 shows the overview of our proposed method that related to each chapter in this dissertation. This research assumes that blind people can reach at the bus stop by themselves. Then, the blind can turn on and take a photo while waiting for the bus. Moreover, three different situations will be considered such as non-congested traffic, congested traffic and non-congested traffic with obstacle consideration in chapter 4, chapter 5 and chapter 6, respectively.

This page intentionally left blank

2. Previous Studies

This chapter discusses related previous studies, the concerning problems of the previous studies, and proposals to mediate the limitations therein.

2.1 Related Previous Research

2.1.1 Public Transportation and Assistive System for Blind People

Since, the majority of blind people and visually impaired people is not able to drive the vehicles by themselves; the public transportation is an important option for blind people who want to travel independently to work, school and shopping. There still are many barriers for blind individuals accessing the public transportation system. Such the print formats of route and timetable, ticket vending machines that have only the touch screen system, and safety warning via visibly displayed, the assistive system does not provide for them [20-22].

The blind people usually choose two main public transportation services for their travel, which consist of railway and bus services. Many useful assistive systems and devices have been proposed for blind people in term of railway and bus services as shown in 2017, Janusz P. *et al.* [23] concerned the tactile graphics at the railway station, and discussed about crucial information, sizes and touch paths on tactile graphic. Moreover, the tactile graphic was used in several countries such as Japan, Korea and Poland. The public buses transportation service is another comfortable choice for blind because there are many bus stops in the city along the public road. In 2007, the Royal National Institution for the Blind (RNIB) of UK [24] developed the talking bus stops in Brighton and Hove city for helping blind people. The talking bus stop was a react unit that provided the real time information for blind users. In addition, AudioTransantiago was an application on portable PC [25] that allowed users to plan the trip and provided the useful information via synthesized voice. Moreover, the smartphone become the useful device for blind people nowadays because there are many developed applications for blind people containing on smartphones.

However, our research focuses on the application for aiding blind people waiting for the bus. Therefore, we would like to discuss intensively about the bus identification and its problems of existing researches.

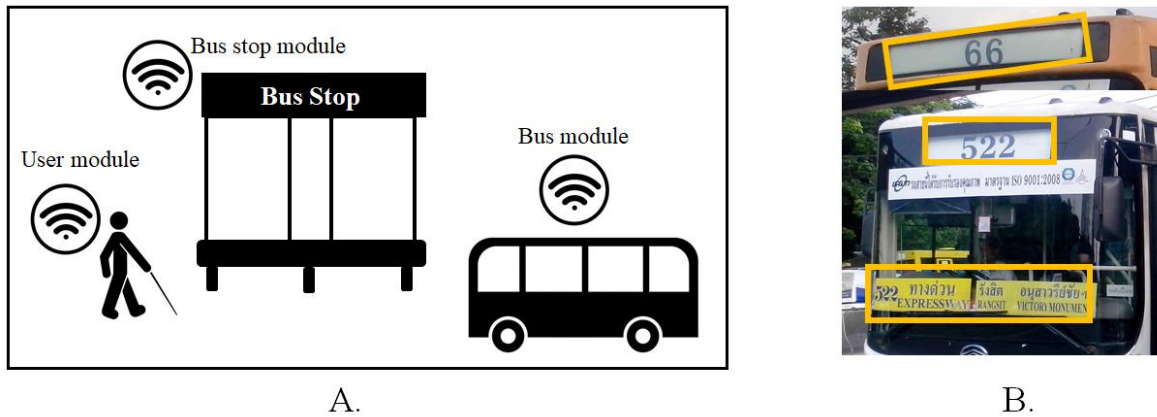


Figure 2.1 Example of two main bus identification system for blind people.

A. Bus identification using the transceiver communication, B. Bus number detection and recognition using image-processing technique.

2.1.2 Bus Identification System for Blind People

The public bus service is the best choice for blind people in daily life because it is easy to access the bus stops in their city and low cost. However, the public bus service still has many problems for blind people who travels independently. Especially, the underdeveloped and some developing countries did not provide the assistive systems for disable people including blind individuals. Therefore, there are many researchers from various countries presenting the bus identification system for blind people.

Although, the bus identification system for blind people has been proposed by many previous researchers, those existing bus identification methods for blind individuals can be categorized into two main systems as shown in Figure 2.1. First system was proposed by transceiver communication use, and second system applied on the image processing field.

2.1.2.1 Transceiver Communication Use for Bus Identification System

The first system uses transceiver communication [8-14] such as RFID, Bluetooth, WiFi or Global Positioning System (GPS). The general concept of this technique is the installation of a receiver-transmitter with the user's device, such as Personal Digital Assistant (PDA) [9-10] or smartphone [8, 13-14]. For example, E. A. B. Santos proposed [8] the wireless interactive system in which the frequency of 2.4 GHz. was used for communication via each installed module between the user, the bus station, and the bus. Users can select the bus line on the smartphone and then the bus station module will receive the requirement from a user module when checking if the bus line is correct. After, the module at the bus station will send the data to the bus module to notify the bus drivers on the bus panel. A similar bus identification system was presented using RFID module [9, 11-12]. Moreover, the VIABUS [14] is a smartphone application in Thailand that applies the GPS module to search the arrival bus. However, it can only search the buses that had the GPS module installed. Although a transceiver communication system is a good idea to identify the bus for visually impaired people, this communication system is quite large for practical use because transceiver modules have to be installed for every bus station, buses, and user. Further, system maintenance is difficult when some part of the module is not working.

2.1.2.2 Image Processing Field for Bus Identification System

On the other hand, the second system proposed is based on an image processing technique [15-19] to capture the bus numbers of the oncoming buses. Image processing is a fascinating technique that can be widely used for several applications, including bus identification for visually impaired people and blind individuals. For example, Pitchakorn Wongta *et al.* [15], in 2016, proposed an automatic bus route number recognition system in Thailand. Their experimental results showed an accuracy of 73.47% for digit segmentation that the researchers themselves collected, that is, the Thai bus dataset. Furthermore, they interviewed some Thai individuals with poor vision concerning the difficulty they faced when traveling independently using public transportation. In the interview, they mentioned reading the bus number is still challenging for them, although they can see a blurry oncoming

bus. The problem of bus number recognition for blind individuals has not only drawn the attention of Thai researchers, but also has also piqued the interest of several researchers from other countries such as Taiwan, Korea, Italy, and the U.S. [16-18]. For example, Ching-Ching Cheng *et al.* [16], in Taiwan, presented a technique for bus route number detection comprising three main processes, namely moving object detection, bus panel extraction, and bus route number detection. Significantly, this resulted in a 100% detection rate. Similar to the approach adopted by Dongjin Lee *et al.* [17], in Korea, they provided a system for automatic number recognition of the bus from natural scenes. Moreover, a Text-to-Speech conversion was conveyed via headphones to the visually impaired users.

Based on above-mentioned system, the bus identification system for blind people is indeed important. Many researchers tried to solve this problem by various method depending on situation of their countries. Our research focus on the bus identification system for blind people in the Thailand and other countries, where have no good system of the public-bus to support the blind people as following below conditions;

1. Arrival time for the oncoming bus cannot be estimated.
2. There is no route bus number announcement to users.
3. Users have to observe the bus number by themselves at the bus stop, and then wave their hand to signal the bus driver to pick them up.

This research related to image processing technique, especially smartphone application.

2.2 Problem of Existing Previous Researches for Bus Identification System

According to existing previous researches of bus identification system, especially the image processing technique. Figure 1.1 shows the other existing proposed methods in the step 2, which consisted of four general processes such as bus detection, bus façade detection, bus number detection and bus number recognition. Although, existing methods seem to be the practical idea, those methods still have many problems when the blind users hold their smartphone freely.

Based on step 2 in Figure 1.1, the blind users have to hold smartphone and take the image freely while waiting for the bus; they cannot know exactly viewpoints of images. For

example, if they are standing behind an obstacle, such as an electricity pole or other big obstacles on the roadside. In this case, they may obtain an image of the whole or part of the obstacle instead of the oncoming bus. Even in ideal situations without obstacles along the road, it is still difficult to inform users how suitable or unsuitable their viewpoints are since it can be impossible to capture the bus on every frame of their video. Moreover, there are various situations of traffic on the road such as congested and non-congested situation, which are uncontrollable conditions as shown in Figure 2.2. For these reasons, other existing methods could not perform the system completely.

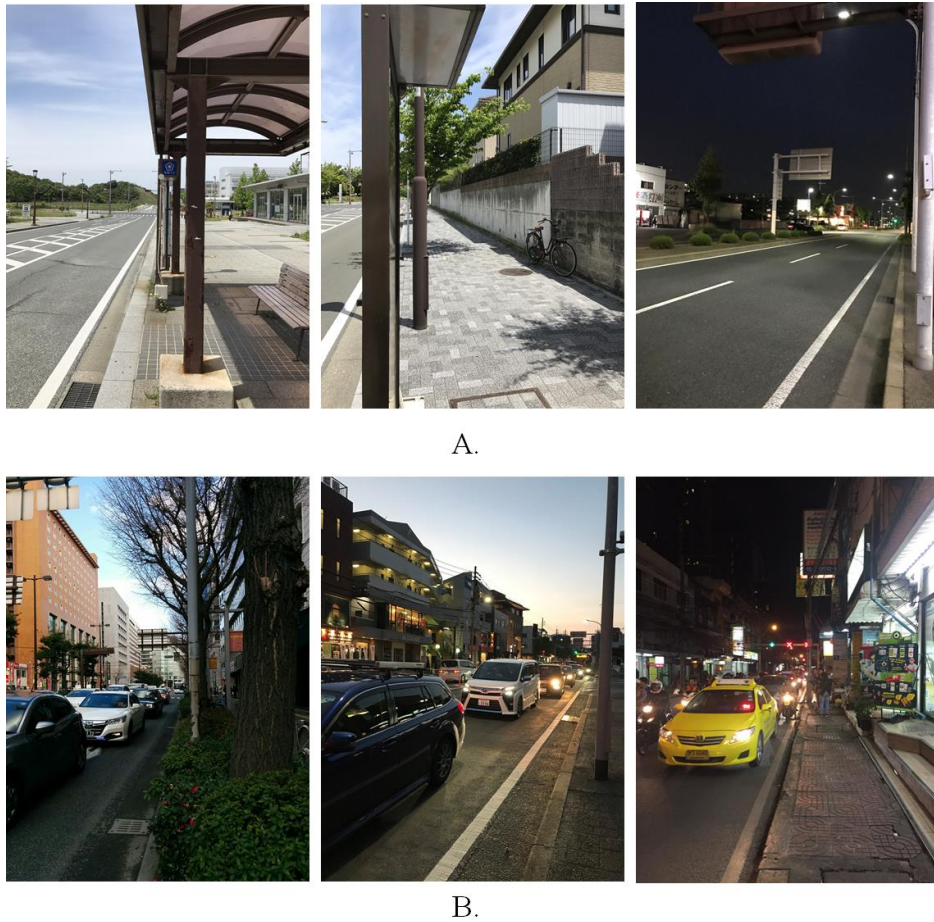


Figure 2.2 Example of image viewpoints for the real situation.

A. Non-congested traffic, B. Congested traffic.

2.3 The Proposed Method

To fulfill the existing system of bus identification for blind people, our study proposes the viewpoint consideration and classification of waiting for the bus. There are three main conditions for our proposed method depending on the traffic situation and obstacle appearance along the road. First condition is the non-congested traffic without obstacle consideration. Subsequently, the situation of congested traffic will be concerned without the obstacle appearance. For third proposed, the obstacles along the road will be focused. Moreover, this study proposes a definition of suitable viewpoint of waiting for the bus that will be explained in chapter 3.

2.3.1 Proposed Method for Non-Congested Traffic

First, the situation of non-congested traffic without obstacle appearance along the road. For this case, we proposed the classification of viewpoint for bus-waiting [26] between suitable and unsuitable viewpoint for various conditions of daytime and nighttime.

Four features are extracted from the road area and vanishing point of perspective image, which related to the suitable viewpoint definition. For this proposed, the combined method of image processing such as Rotational Invariant of Uniform Local Binary Pattern ($LBP_{P,R}^{riu2}$) and Hough Line Transformation are used.

2.3.2 Proposed Method for Congested Traffic

Second, we focused on the congested traffic situation without obstacle appearance, and classified the viewpoint for daytime and nighttime [27]. For this situation, the distribution of cars on the road will be calculated due to the road area could not use for feature extraction.

The YOLO technique [51] is applied for car detection, and each detected car is plotted in x, y coordinate. Then, nineteen features of car distribution are considered such as number of cars, R^2 , slope and y- interception of linear regression. Finally, our study optimizes the accuracy of classification process using several classifiers and features.

2.3.3 Proposed Method for Obstacle Detection along the Road

Third, the situation of obstacle appearance along the road is concerned [28]. There are two main steps for this proposed method consisting of road area detection and obstacle detection. For the road area detection step, we applied the similar technique as mentioned briefly in section 2.3.1, but there are additional processes and differences for obstacle detection process.

Since, some obstacles along the road do not affect to bus number detection and recognition, therefore the system have to classify and estimate the height of obstacle that related to the suitable viewpoint definition. For the obstacle detection process, we separate the method into two sub-process; first is to find the obstacle's position along the road using the vertical projection process and other existing techniques; and second process is the obstacle's high estimation using the color moment calculation.

This page intentionally left blank

3. Suitable Viewpoint Definition of Waiting for the Bus

3.1 The Possible Image Viewpoints of Waiting for the Bus Using Smartphone

Our system was designed, as shown in Figure 1.1, which consisted of four main steps of overview. Especially, the blind users have to hold the smartphone independently to take the photo while waiting for bus. There are various possibilities of image viewpoints appearance depending on the user holding their smartphone randomly. For this study, the driving on the left side of the road will be considered such as Thailand and Japan. In addition, the possible image viewpoints of waiting for the bus using smartphone are separated into five main viewpoints as shown in Figure 3.1. Although, some viewpoints may show other patterns excluding our mentioned for five main viewpoints, but those viewpoints occur from combination of five main viewpoints. For example, Figure 3.2A. shows the combination of excessive high-angle and excessive left direction of viewpoint. Moreover, Figure 3.2B. shows the excessive low-angle of viewpoint with excessive left direction. On the other hand, the viewpoint of too much for low-angle and right direction is shown in Figure 3.2C.



Figure 3.1 Five main possible image viewpoints of waiting for the bus.

A. Excessive high-angle of viewpoint, B. Excessive low-angle of viewpoint, C. Excessive right direction of viewpoint, D. Excessive left of viewpoint, E. Suitable angle and direction of viewpoint.

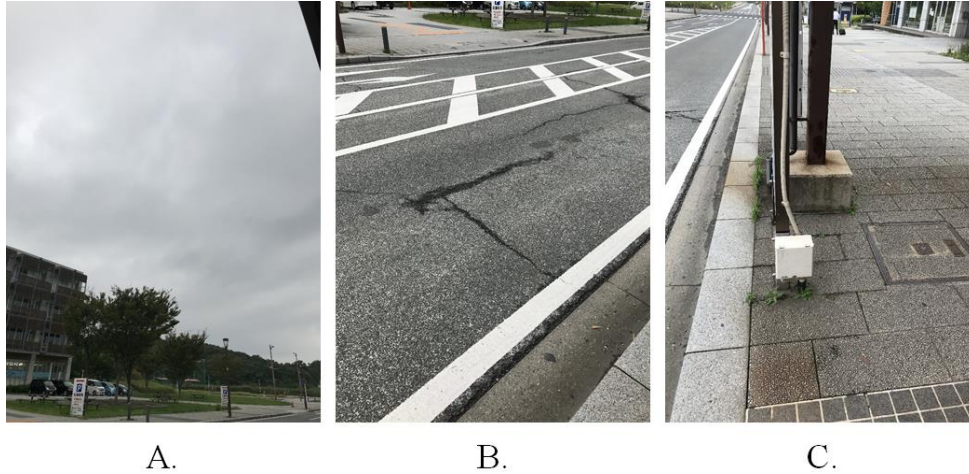


Figure 3.2 Example of other image viewpoints.

A. Excessive high-angle and left direction, B. Excessive low-angle and left direction, C. Excessive low-angle and right direction.

3.1.1 Excessive High-Angle of Viewpoint

An excessive high-angle of viewpoint is one of possible image viewpoint when the blind users hold the smartphone freely. This viewpoint pattern will be appeared when users tilt their smartphone with high-angle or excessive high tilt of smartphone. For example, the excessive high-angle of viewpoint in Figure 3.1A. shows the large area of sky and roof of bus-stop in the image, and small area of the road. Some cases of excessive high-angle viewpoint may not appear the road area or any cars on the road, which means that the bus number will not appear by this image viewpoint pattern.

3.1.2 Excessive Low-Angle of Viewpoint

A pattern of excessive low-angle of viewpoint will be appeared when users hold the smartphone with low-angle or excessive low tilt, which is opposite angle of excessive high-angle of viewpoint. Figure 3.1B. shows an example of excessive low-angle pattern that the huge road area is captured. For this case, it is also difficult to detect and recognize the bus number of oncoming bus because the angle of smartphone is in quite low but the bus number is in the high position of image. Therefore, this viewpoint is the unsuitable for bus-waiting.

3.1.3 Excessive Right Direction of Viewpoint

Although, users hold their smartphone for suitable angle (not excessive high or low), the image viewpoint still might be unsuitable direction of waiting for the bus. Figure 3.1C. is another pattern of unsuitable viewpoint when the users hold their smartphone in right direction excessively or excessive right panning of smartphone. For example, the Figure 3.1C. appears the large area of footpath on the right side of image but narrow area of road on the opposite side. Even, the road area will be appeared in the image for this example, but some cases will not show the road area when the users excessively pan their smartphone to the right direction.

3.1.4 Excessive Left of Viewpoint

An excessive left of viewpoint is one of possible pattern when the blind people hold the smartphone to the left direction extremely. This image pattern usually appears the large area of road that is very close to the user standing point as shown in Figure 3.1D. Although, it seems possible to detect the bus number for this case, but it is too close for the user. Since, the users who waiting for bus, in Thailand, have to observe the bus number by themselves at the bus stop, and then wave their hand to signal the bus driver to pick them up. Therefore, the time for users to call the bus is not sufficient in this viewpoint pattern.

3.1.5 Suitable Angle and Direction of Viewpoint

The possible viewpoint for bus number detection and recognition will be shown when the user hold the smartphone appropriately. Figure 3.1E. shows the suitable viewpoint of waiting for bus, which this viewpoint is possible to detect the bus number of oncoming bus, and there is sufficient time for users to decide and make the signal to the bus driver.

Although, the suitable viewpoint is easy to distinguish for human vision and brain processing, but the exactly definition of suitable viewpoint of waiting for the bus is crucial for a computer vision term. Moreover, there is no previous definition for this situation (excepting our public article in [26]). Therefore, this study will discuss and explain the suitable viewpoint in the section 3.2.

3.2 A Suitable Viewpoint Definition of Waiting for the Bus using Smartphone

There is no previous study that proposing the viewpoint classification for blind individuals waiting at a roadside bus stop. Our study discusses and defines the suitable viewpoint of waiting for the bus based on smartphone use in this section. According to above-mentioned involving the possible image viewpoint of bus-waiting in section 3.1, there are two main factors of smartphone holding affecting to those image viewpoints appearance. Figure 3.3 shows the main factor of suitable viewpoint, which consists of tilt of camera and panning of camera.

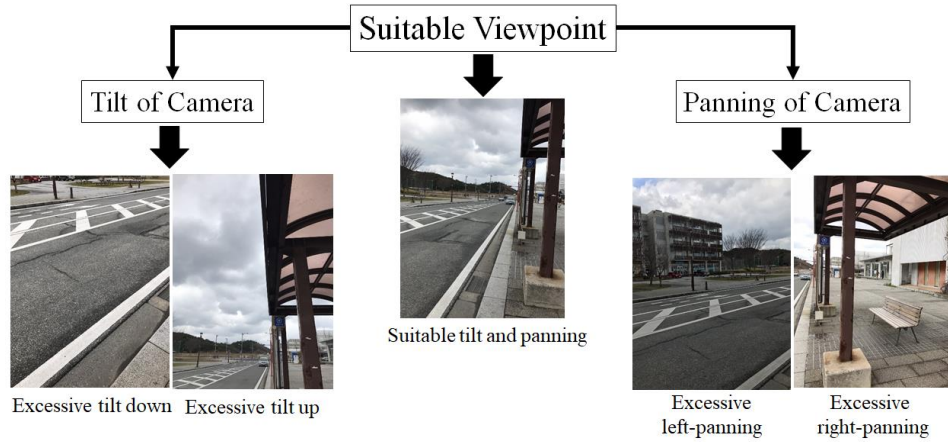


Figure 3.3 Two main factors of suitable viewpoint appearance while waiting for the bus.

3.2.1 Suitable Tilt Definition of Smartphone Use while Waiting for the Bus

Tilting is a cinematographic technique, which the camera stays in a fixed position but rotates up or down in vertical plane [29]. The result of viewpoint is similar to someone raising or lowering the head up or down for looking some things.

Tilt of smartphone is one of factor for suitable and unsuitable viewpoints. Tilt of smartphone is shown when the users do not hold the smartphone in vertical axis or it is not perpendicular with X-axis. There are two patterns of tilt such as tilt up and tilt down as shown in Figure 3.4B and Figure 3.4C respectively. However, the tilt of smartphone excessively let the viewpoints to be unsuitable for bus-waiting. Therefore, the definition of suitable tilt of smartphone while waiting for the bus is *holding the smartphone vertically* as shown in Figure 3.4A.

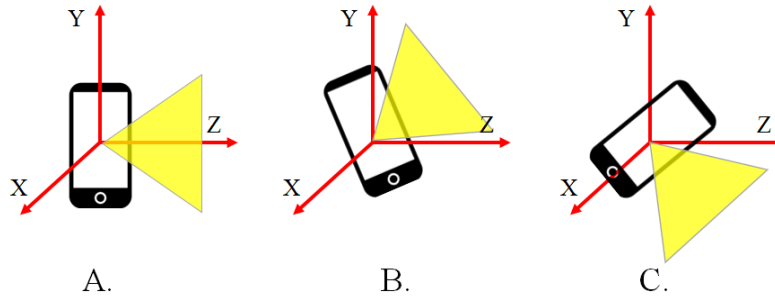


Figure 3.4 Holding smartphone and its tilt.

A. Holding smartphone in vertical axis (non-tilt), B. Holding smartphone in tilt-up pattern, C. Holding smartphone in tilt-down pattern.

Holding the smartphone vertically means that the accelerometer value provides ideally the x, y, and z as 0.0 m/s^2 , 10.0 m/s^2 , and 0.0 m/s^2 , respectively. However, the definition of suitable tilt with holding the smartphone vertically is extremely fixed and impractical for users. This study sets the range of criteria for suitable tilt of tilt-down and tilt-up that is possible to detect and recognize the bus number. The suitable of tilt-down is defined by $y = 9.5 \text{ m/s}^2$ and $z = 3.5 \text{ m/s}^2$, as shown in Figure 3.5. Moreover, the suitable tilt-up, in Figure 3.6 m/s^2 , is set by $y = 9.0 \text{ m/s}^2$ and $z = -4.0 \text{ m/s}^2$. The x-value of accelerometer is ignored, because tilt-down and tilt-up of smartphone do not change the x-value significantly, as shown in Figure 3.7.



Figure 3.5 Suitable of tilt-down ($y = 9.5 \text{ m/s}^2$ and $z = 3.5 \text{ m/s}^2$ of accelerometer value).



Figure 3.6 Suitable of tilt-up ($y = 9.0 \text{ m/s}^2$ and $z = -4.0 \text{ m/s}^2$ of accelerometer value).

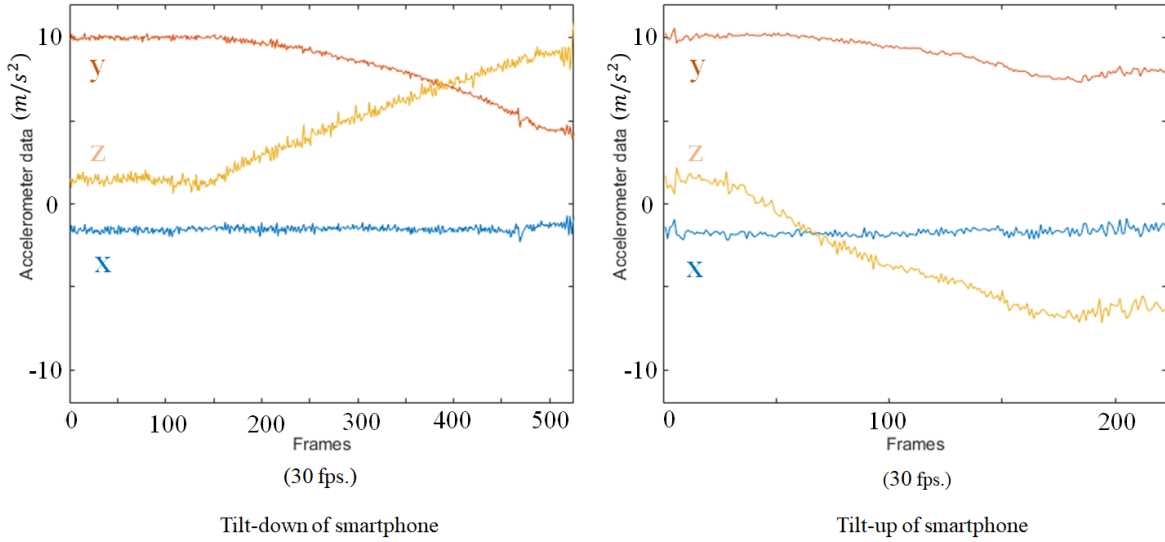


Figure 3.7 Tilt-down and tilt-up of smartphone with x, y and z of accelerometer values.

3.2.2 Suitable Panning Definition of Smartphone Use while Waiting for the Bus

Panning, in term of photography and cinematography, is swiveling a still or camera horizontally from fixed position [30]. It seems to the motion of someone who turn their head on their neck from left to right or right to left as shown in Figure 3.8.

Various viewpoints can be appeared by panning smartphone, which the interpretation and definition of suitable panning of smartphone while waiting for the bus is complex than the suitable tilt. Since, the suitable viewpoint for our application involves the route bus number recognition; however, the real scenarios of waiting for bus can occur many

viewpoints depending on the situations on the road. For example, the non-congested traffic situation means having few cars or no car on the road including the bus. Thus, the suitable viewpoint of smartphone panning must be defined, although, there are no cars and buses in the viewpoints.

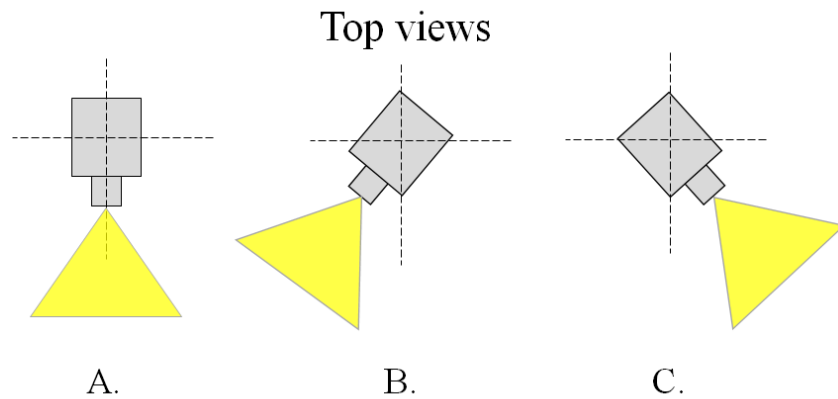


Figure 3.8 Camera panning.

A. Direct panning, B. Right panning, C. Left panning.

3.2.2.1 Assumption of Waiting for the Bus and Our System Configuration

This study focuses on the public bus service in Thailand, which the vehicle drives on the left side of the road as same as Japan system. Moreover, our system involves the route bus number detection and recognition for both of daytime and nighttime. In order to interpret the suitable panning and longest distance for route bus number recognition of various situations. Therefore, our application configuration will consider the conditions of environment of waiting for the bus such as the route bus number size in Thailand, all characters appearing on the public bus in Thailand, and illumination in day and night times.

1. Route Bus Number Size in Thailand

In Thailand, there are many sizes of bus number depending on the type of buses as shown in Figure 3.9. However, the standard size of route bus number in Thailand is set as 15×15 cm. Thus, this study simulated the board as same as the standard size of Thai-public bus number for estimating the longest distance of bus number recognition as shown in Figure 3.10.



Figure 3.9 Different types of Thai-public bus and bus number size.

([31], [32] refer to Figure 3.9A. and Figure 3.9B., and [33] refers to Figure 3.9C. and Figure 3.9D.)

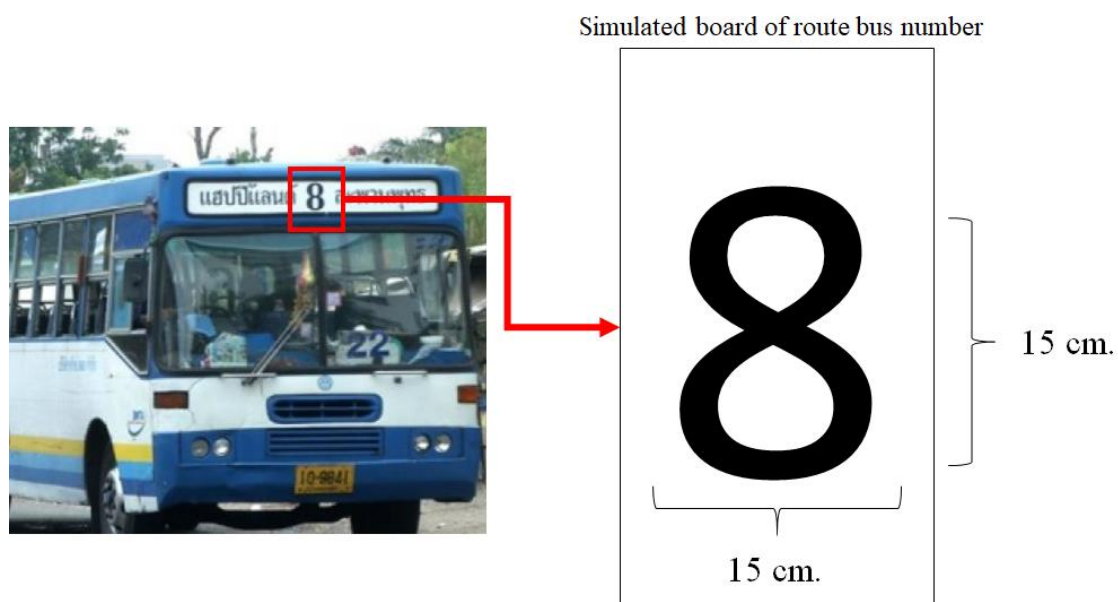


Figure 3.10 Simulated route bus number in Thailand by 15 x 15 board size.

2. All Characters Appearing on the Public Bus in Thailand

According to the Bangkok Mass Transit Authority (BMTA) [34-35], there are 108 routes for the public bus and 3,509 of buses for daily service in Bangkok metropolitan region. Based on our observation from BMTA website, there are the numbers, some English and Thai alphabets appearing on the route bus number as shown in Table 3.1.

The necessary characters of Thai bus number consists of the number (0-9), English alphabets (A and B), and Thai alphabets (ก ต ป อ). Therefore, these characters will be tested for estimating the longest distance to recognition of our study.

Table 3.1 Characters appearing on the Thai-route bus number.

Types	Characters	Example of route bus number
Number	0 - 9	1, 2, 3, 4, and 5
English Alphabets	A and B	A, A1, B, and B1
Thai Alphabets	ก ต ป อ	7ก, ต10, and ปอ84

3. Illumination in Day and Night Times

In order to provide the system for using in day and night times, this study will be discussed for this condition. Illumination is one of factor that is very various and uncontrollable, which can be changed depending on the conditions of sunlight, weather, and lighting of environment. Especially, our application will be used for the outdoor scenario in both of day and night times. Moreover, the illumination can effect to the vision for bus number detection and recognition distance. Unfortunately, to estimate the longest distance for bus number recognition for all real illuminations in both of day and night times, outdoor scene, is impossible, because of above reasons. However, this study sets the ideal condition for daylight and nightlight, which we measure the LUX range using digital LUX meter (LX-1332B), in Figure 3.11, as shown the important specification of equipment [36] in Table 3.2. Based on our measurement of illumination, the daylight and nightlight conditions were set as 210.00-230.00 lux and 0.80-1.00 lux respectively.



Figure 3.11 Digital LUX meter for our study.

Table 3.2 Specification of digital LUX meter (LX-1332B).

Specification	Range and Value
Display	3.5 digits
Range of measurement	200 Lux, 2,000 Lux, 20,000 Lux, and 200,000 Lux.
Accuracy	$\leq 10,000$ Lux: $\pm 4\%$, $\geq 10,000$ Lux: $\pm 5\%$
Repeatability	$\pm 2\%$
Temperature characteristic	$\pm 0.1\%$ / $^{\circ}\text{C}$

3.2.2.2 Estimation of Longest Distance for Route Bus Number Recognition

In order estimate the longest distance for route bus number recognition, this study set the environment according to mentioned in section 3.2.2.1. This section explains and shows the longest distance for route bus number recognition for daylight and nightlight conditions.

Since, the longest distance for route bus number recognition is related by many factors such as illumination, size of objects and resolution of image. However, the illumination of daylight and nightlight were defined by 210.00-230.00 lux and 0.80-1.00 lux respectively, using lux meter. Moreover, the size of route bus number was set as 15×15 cm

as same as the standard size of Thai bus number. In term of image resolution, this study selected the 800×600 px that is suitable for our application.

1. Defining the Distance for Testing

After setting the environment illumination as mentioned above, this study defines the different distances for estimating the possible longest distance of route bus number recognition. For the experiment, the smartphone camera was used by iPhone7 for taking the image, which the image resolution was 800×600 px. Moreover, there are three different distances for this testing including 10 m, 15 m, and 20 m.

Figure 3.12 shows the example of image output by capturing in different distances. At the 10 m, the image output shows clearly image of number 8. However, the clarity of image output will be decreased when the capturing distance is increasing.

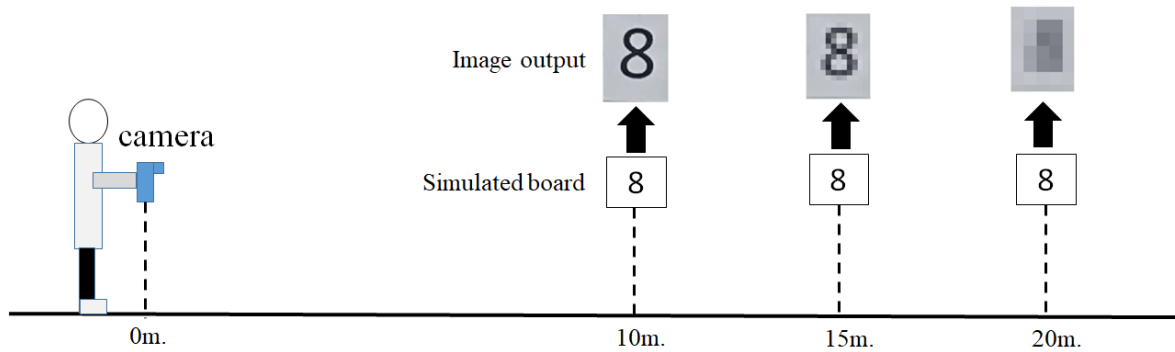


Figure 3.12 Example of different capturing distance and its image output.

2. Testing the Possible Longest Distance for Route Bus Number Recognition

All characters of route bus numbers as mentioned in section 3.2.2.1 will be considered using the optical character recognition (OCR) [37-39]. The OCR technique is widely used for text recognition that it can translate the images of typewritten or hand written characters into machine editable format [40]. There are three main processes for text recognition using OCR. First step is to select the region of interest (ROI) in an image, then, the ROI area is converted to binary image. Finally, the binary output image is processed by OCR technique as shown in Figure 3.13.

Capturing simulated bus number

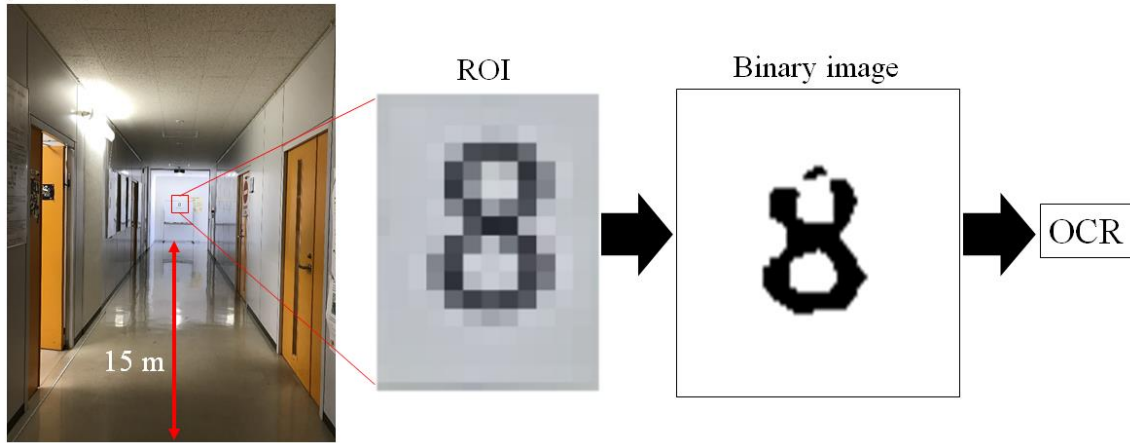


Figure 3.13 Processing step of OCR testing by capturing the simulated bus number.

Table 3.3 Possibility of route bus number recognition using OCR technique.

Characters	Distance (meter)		
	10	15	20
0	✓	✓	✓
1	✓	✓	✓
2	✓	✓	✓
3	✓	✓	✓
4	✓	✓	✓
5	✓	✓	✓
6	✓	✓	✓
7	✓	✓	✓
8	✓	✓	✓
9	✓	✓	✓
A	✓	✓	✗
B	✓	✓	✗
ก	✓	✓	✗
ข	✓	✓	✗
ค	✓	✓	✗
อ	✓	✓	✗

✓ the character can be recognized

✗ the character can not be recognized

Three different distances of capturing will be tested for daylight and nightlight at 10 m, 15 m, and 20 m respectively. Table 3.3 shows the possibility of route bus number recognition using OCR for different distances. According to our testing, thirteen images for all characters of necessary Thai bus number could perform at 10 m and 15 m. However, it could not provide the good performance at the 20 m because the English and Thai alphabets provide the wrong recognition. Therefore, this study chose the possible longest distance for route bus number recognition at 15 meters (reference distance).

3.2.2.3 Suitable Panning while Waiting for the Bus

Up to now, the possible longest distance for route bus number recognition has already decided by 15 meters. This section will define the suitable panning of smartphone's camera while waiting for the bus, although there are no any cars or buses on the road.

The suitable panning definition is quite complex, but it can be estimated by the longest distance for route bus number recognition. Firstly, capturing a photo of the oncoming bus at the bus stop within the 15 m, as shown in Figure 3.14A. The distance is measured from the standing point of the user to the façade of the oncoming bus. In order to estimate the image viewpoint without the oncoming bus, the vanishing point (V_p) of the perspective image is defined by two convergent lines of the white line marker on the edge of the road. After, an L horizontal line is drawn from V_p to the square boundary of the detected bus, as shown in Figure 3.14A. L line is measured as 8.70% (0.49 inch) compared with the width of the original size of the image (5.63 in). In case there is no bus on the road, the L line can be used to estimate the bus's position at the longest distance for route bus number recognition. For example, Figure 3.14B shows the suitable camera panning. By marking the vanishing point and drawing the L line, the system can estimate the position of the bus in the image where the route bus number will appear when the bus arrives at that position. On the other hand, when the camera pans lightly, as shown in Figure 3.14C, the bus position can be estimated, but appears with just half of its façade.

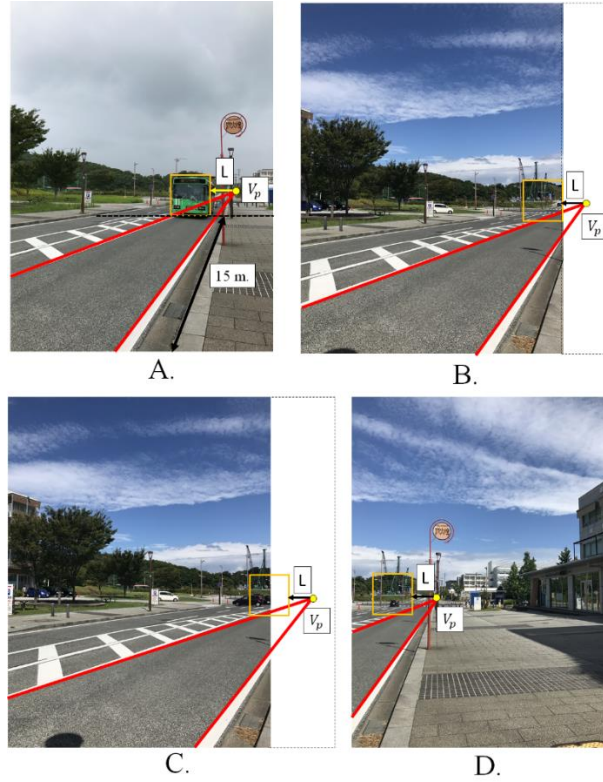


Figure 3.14 Estimation of viewpoints with/without oncoming bus.

A. Oncoming bus at a distance of 15 meters from the camera, B., C., and D. Viewpoint estimation without oncoming bus using vanishing point and L line.

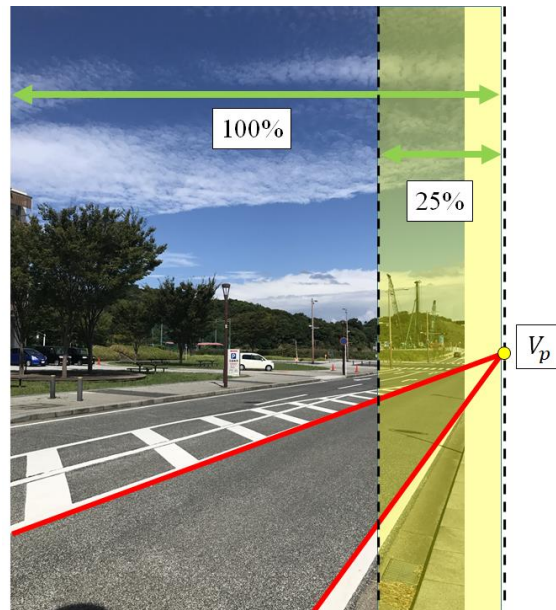


Figure 3.15 A suitable viewpoint with 25% of parameter setting.

Thus, it can be claimed that the image viewpoint is unsuitable for the bus number recognition because the route bus number cannot be recognized at the reference distance. Although the bus number can be recognized in Figure 3.14D at the reference distance, the moving bus will be seen in the image frame only for a short time. Therefore, it would be better if the camera were panned to the left side in order to see the bus for a longer period of time. Actually, the image viewpoint in Figure 3.14B is the *perfect camera panning* to recognize the route number of the bus because the bus number can be recognized at the farthest right side of image; this is the first position in the image viewpoint to see the bus. However, it will be extremely impractical to define only a single suitable viewpoint for the users to hold their smartphone. Consequently, this study defines *the suitable panning of viewpoint within 25% of the range between V_p and the farthest left side in the perfect camera direction that the vanishing point of perspective images have to fall in*, as shown in Figure 3.15. However, setting a value of 25% is an assumption of the proposed application.

3.3 Summary

The suitable viewpoint while waiting for the bus is defined by two main factors such as tilt and panning of smartphone. Tilt of smartphone is a factor for suitable viewpoint appearance, which the suitable tilting is defined by *holding the smartphone vertically*. The suitable tilt-down of smartphone was set by $y = 9.5 \text{ m/s}^2$ and $z = 3.5 \text{ m/s}^2$. Moreover, the tilt-up of smartphone was defined by $y = 9.0 \text{ m/s}^2$ and $z = -4.0 \text{ m/s}^2$. For the suitable panning definition, it is quite complex comparing to suitable tilt situation. However, the suitable panning definition is explained that *the suitable viewpoint within 25% of the range between V_p and the farthest left side in the perfect camera direction that the vanishing point of perspective images have to fall in*.

This page intentionally left blank

4. Classification of Viewpoints while Waiting for the Bus in Situation of Non-Congested Traffic

This chapter includes four main sections. First, the objective of this study will be explained. Second, the proposed method for classification of viewpoint in case of non-congested traffic such as the Rotational Invariant of Uniform Local Binary Pattern ($LBP_{P,R}^{riu2}$), k -means clustering, Hough Line Transformation, and others. Third, the experiments and results of our proposed method is shown by the percentage of accuracy. Finally, the result and the proposed methods will be discussed in the last section of this chapter.



Figure 4.1 Different patterns of road area with different illuminations between daytime and nighttime.

4.1 Objective of Classification of Viewpoints in case of Non-Congested Traffic

Although, there are various situations appearing on the road, this study will first focus on the non-congested traffic situation. The main objective of this study is to classify the viewpoints while waiting for the bus, in case of non-congested traffic. To find the suitable panning of smartphone, which related to the suitable viewpoint, this study assumed that the

users have held their smartphone vertically. Moreover, the proposed concept for interpreting viewpoints is to extract essential features from the road area via images. These features comprise road length, vanishing point, and the percentage of road area. However, there are many patterns of road area in the images depending on different viewpoints and illumination as shown in Figure 4.1, which this study will be explained, and test under both daytime and nighttime.

4.2 Proposed Method for Classification of Viewpoint based on Non-Congested Traffic

Since, the objective of this study is to classify the viewpoint while waiting the bus based on the road area segmentation. The difficulty of outdoor scenes segmentation, especially road area, is uncertain illumination and unstructured shapes. Hence, this study presents the new combination method for road segmentation and useful features extraction for image viewpoint classification, as shown in Figure 4.2. Four main processes consist of finding the reference of road area, pre-viewpoint classification, road segmentation and feature extraction, and viewpoint classification.

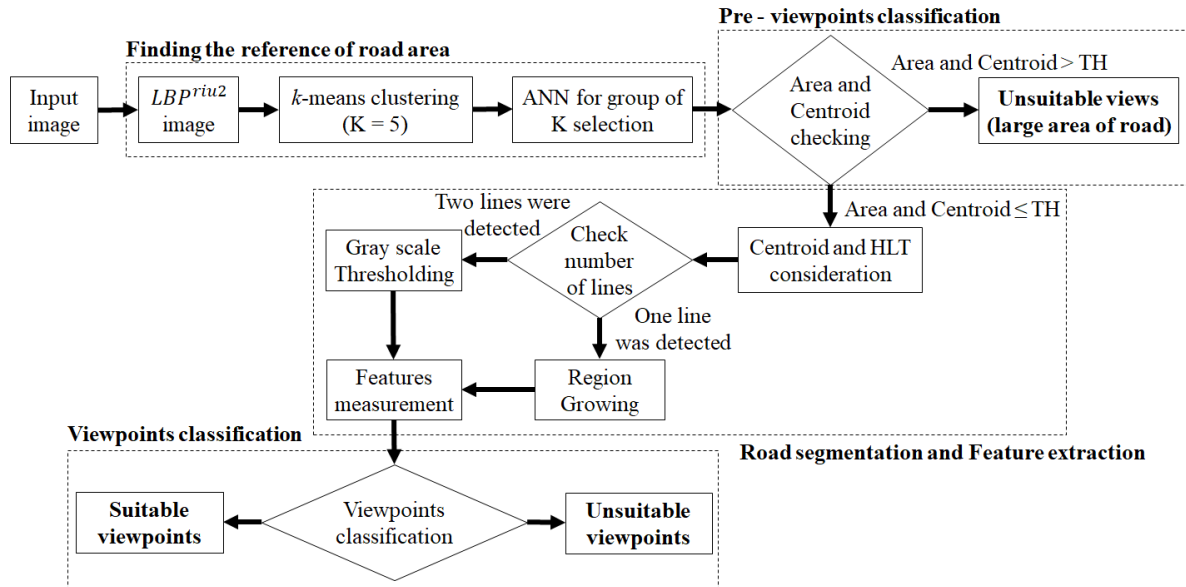


Figure 4.2 Block diagram of the proposed method.

4.2.1 Finding the Reference of Road Area

Generally, roads have an unstructured shapes and different illumination for each image viewpoint. Hence, the first step of the proposed method is finding the reference of road area in the images. There are three steps for finding the reference of road area such as the converting image by Rotational Invariant of Uniform Local Binary Pattern ($LBP_{P,R}^{riu2}$), k -means clustering step, and selecting the group of k -means using Artificial Neural Network (ANN).

4.2.1.1 Rotational Invariant of Uniform Local Binary Pattern

Image texture is one useful feature for image segmentation, especially the Rotational Invariant of Uniform Local Binary Pattern ($LBP_{P,R}^{riu2}$), as proposed by T. Ojala *et al.* [41]. $LBP_{P,R}^{riu2}$ is the technique for texture analysis of an images that finds the relation between the gray-scale value of the center (g_c) and its neighbors (g_p) as written as Eq. 4.1. Figure 4.3A shows the pattern of $LBP_{P,R}^{riu2}$ that this proposed method defines, the number of neighbors (P) and radius (R) as 8 and 1 respectively.

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c); & U(LBP_{P,R}) \leq 2 \\ P + 1; & otherwise \end{cases} \quad (4.1)$$

$$U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)| \quad (4.2)$$

Each gray scale value of neighbor (g_p) will be compared with g_c ; then, the value of 0 and 1 will be set for $g_p - g_c < 0$ and $g_p - g_c \geq 0$ respectively. Nevertheless, only uniform LBP ($U \leq 2$) is considered, where $U \leq 2$ is the pattern of LBP that has no transition ($U = 0$) or two transitions ($U = 2$), as shown in Figure 4.3B wherein the black and white points represent 0 and 1, respectively. For example, the binary code of 00000000 is defined by $U = 0$, and 10000111 is represented by $U = 2$. In case of non-uniformed pattern ($U > 2$), it is shown in Figure 4.3C that the binary codes are 11001011 and 10101010 for $U = 4$ and $U = 8$,

respectively. According to uniform pattern limitation, there are $P + 2$ possible output values that can be represented at the center coordinate point.

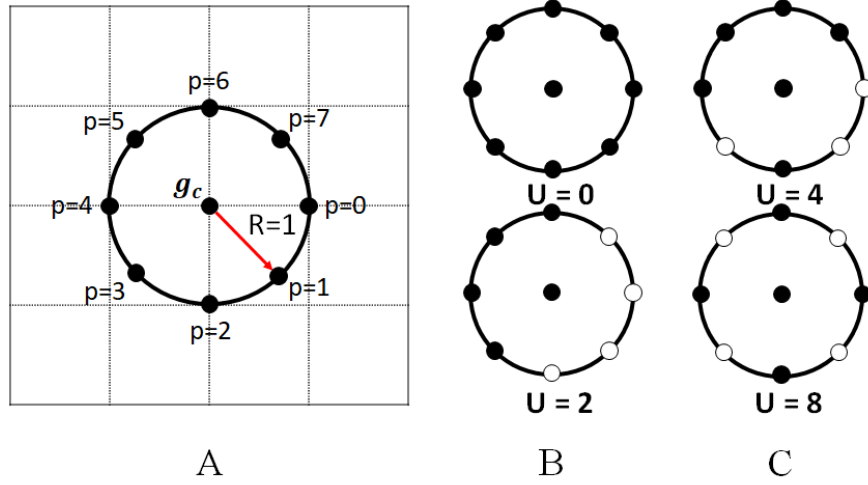


Figure 4.3 $LBP_{P,R}^{riu2}$ with $P = 8$ and $R = 1$.

A. Example structure of $LBP_{8,1}^{riu2}$, B. Uniform pattern of $LBP_{8,1}^{riu2}$ for $U = 0$ and $U = 2$, C. Uniformed pattern of $LBP_{8,1}^{riu2}$ for $U = 4$ and $U = 8$.

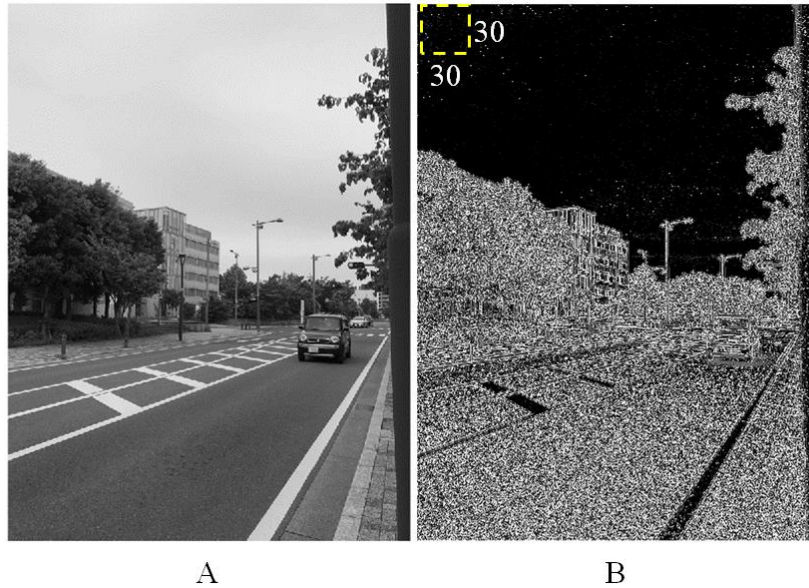


Figure 4.4 Example of transformed image by $LBP_{8,1}^{riu2}$.

A. Gray-scale image, B. $LBP_{8,1}^{riu2}$ transformation image with 30×30 px of sub-window.

This study applies the $LBP_{8,1}^{riu2}$ method on the gray-scale image, as shown in Figure 4.4A. In Figure 4.4B, the complete transformation image of $LBP_{8,1}^{riu2}$, where each pixel is represented by a value between 0 and 9, is shown because the principle of $LBP_{p,R}^{riu2}$ technique will show the uniform pattern ($U \leq 2$) as explained in section 4.2.1.1. Next, a sub-window of 30×30 pixel size is designed in order to calculate the histogram of $LBP_{8,1}^{riu2}$ transformation image for each sliding sub-window.

4.2.1.2 k -means Clustering Process

After each sliding sub-window of $LBP_{8,1}^{riu2}$ transformation image is calculated as shown in Figure 4.5, the k -means clustering will be applied for next step.

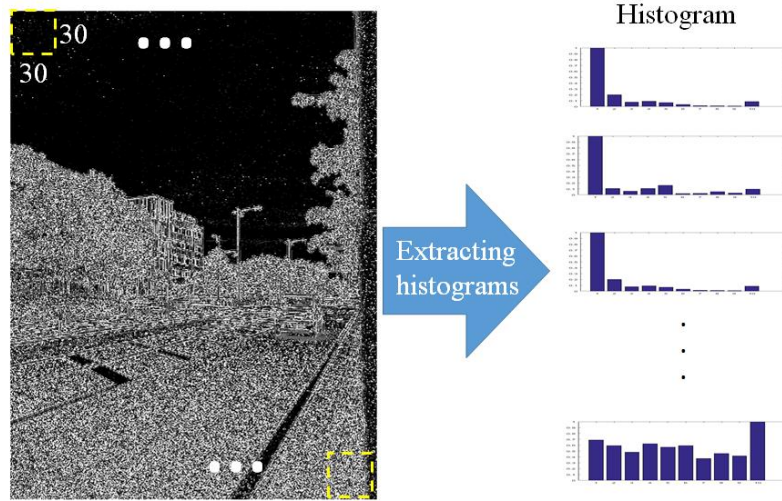


Figure 4.5 Step of histograms extraction for each sliding sub-window.

The k -means clustering method [42] is a widely used for grouping the similar data, which is a type of unsupervised learning. This technique will separate the group of data into K categories that depended on the user defined or applications. Moreover, the centroid value will be calculated for representing each data group. When the new unknown data is assigned to the system, the distance calculation between unknown data and each centroid will be measured. The shortest calculated distance will be selected, and the unknown data will be labeled as same as that centroid. For this study, the k -means clustering defining with $K = 5$

(setting by trial and error) and Euclidean distance measurement. For example, five different labels are separated and shown in Figure 4.6 with the representative histogram of each label (k_n), which the each histogram frequency will be divided by its highest value for normalized frequency.

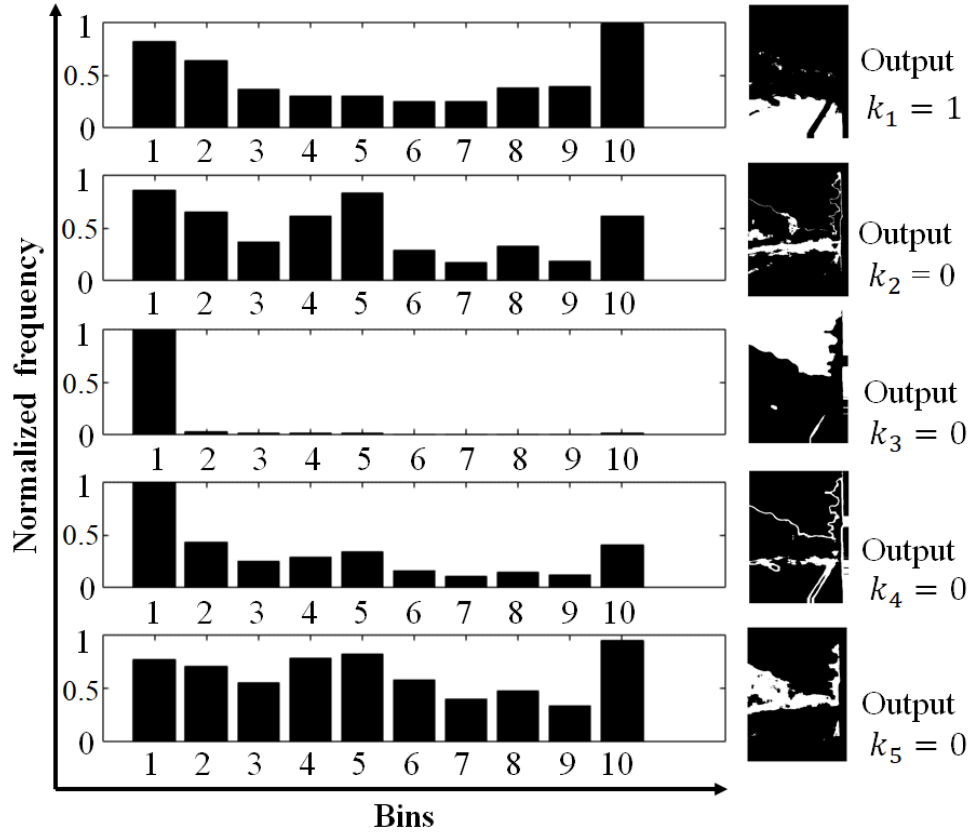


Figure 4.6 Five different $LBP_{8,1}^{riu2}$ histogram representations from k -means.

4.2.1.3 Selecting Group of K using Artificial Neural Network

According to k -means clustering process in section 4.2.1.2, there are five different histogram patterns corresponding to $LBP_{8,1}^{riu2}$ image area. In order to select the group of k -means that contain the road area, this step explains the selection of k_n using Artificial Neural Network (ANN).

ANN duplicated the structure of biological neural that consist of dendrite, cell body and axon for collecting, processing and output signal respectively [43]. Likewise, ANN

consists of input layer (X), hidden layer (Z) and output layer (Y). Each input has different weight values, after that multiply input with weight value. When sum of these values more than the threshold, output to the next neuron as shown in Figure 4.7.

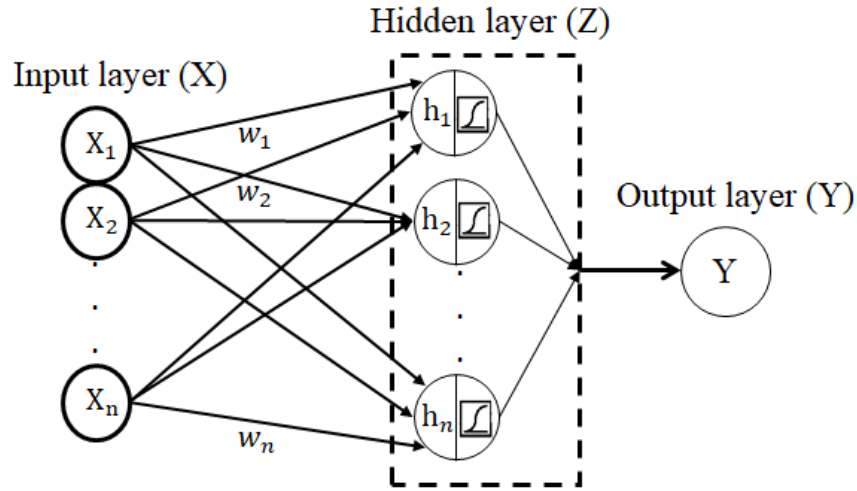


Figure 4.7 Basic structure of Artificial Neural Network.

Although, there are many types of ANN processing, this study selected the feedforward backpropagation of ANN [44]. For the ANN design, the input layer comprises ten nodes for ten bins of histogram. Then, each node of the input layer is multiplied by different weight values and summarized in the hidden layer. Subsequently, the log-sigmoid transfer function is used for obtaining the value between 0 and 1. The output layer defines only one node for arriving at 0 or 1 that indicates whether the label does or does not contain the road area, respectively. For example, in Figure 4.6, the output of k_2, k_3, k_4 , and k_5 show the value as 1, however, there is only one label providing as 1 value as shown in k_1 , and that label contains the road area. After, the group pixel on the left side and the bottom of the image is selected as the *reference road area* because the image viewpoints for bus-waiting always appear in the road area in that part of the images. Furthermore, the performance of ANN selecting k_n will be shown in section 4.3.1.

4.2.2 Pre-Viewpoint Classification

There are many patterns of captured images depending on the panning and tilting of smartphone. One pattern of unsuitable viewpoint will be appeared when the user panning the smartphone to the left hand excessively. Figure 4.8 shows the example of unsuitable viewpoint with the excessive left panning, which this viewpoint shows the too large area of road and close to the user when the bus comes. Therefore, in this section, the viewpoint classification for unsuitable viewpoint, in case of a large area of road, will be explained.



Figure 4.8 The suitable viewpoint with excessive left panning.

Figure 4.9A, one of the unsuitable viewpoints for bus recognition can appear when the area of the road is large. In Figure 4.9A and Figure 4.9B, the difference in binary images that show the reference road area. In order to distinguish the unsuitable viewpoint of excessive left panning, there are two important features for separating the large area of road viewpoint from other views. The first feature is the percentage of reference of the road area because the large area of the road usually takes up more of the frame. Moreover, centroid coordinate point (C_x , C_y), as the red point in Figure 4.9, is calculated for the second feature.

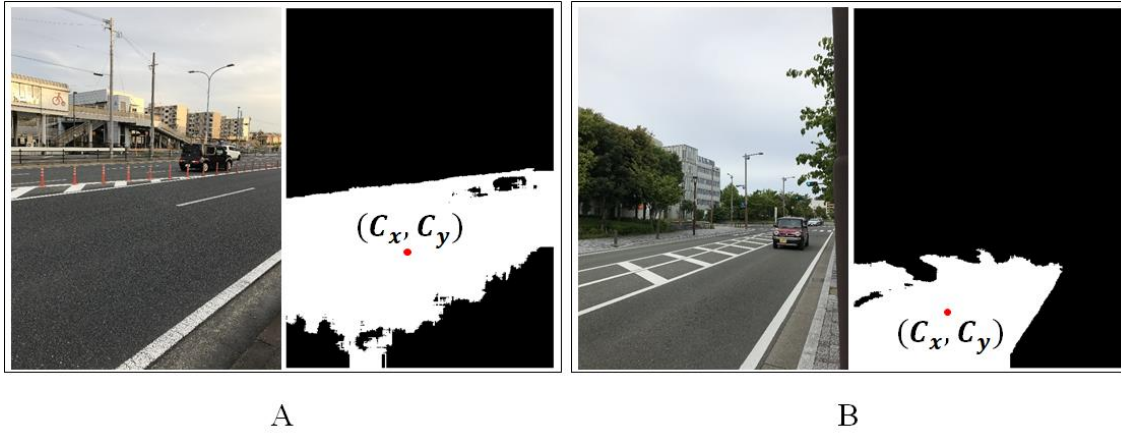


Figure 4.9 Comparison of excessive left panning viewpoint and other viewpoint patterns.

A. Example of unsuitable viewpoint for a very large road area case and its reference road area, B.

Suitable viewpoint and its reference road area.

To set the criteria for large road area viewpoints, this study extracts two features of percentage of road area and percentage of road centroid C_x as shown in E.q. 4.3 and E.q. 4.4. A set of 100 sample images of large road area viewpoints and other views are used for feature consideration. Figure 4.10 shows the criteria setting for the large road area by plotting the extracted features between large road area viewpoints and other viewpoints. The blue dots are the representative features of the percentage of road area and centroid C_x . Obviously, the graph shows the difference between two groups of large road area and other viewpoints (red dots). Consequently, this study sets the criteria of Large Road Area (LRA) case as shown in Eq. 4.5. However, the accuracy of pre-viewpoint classification will be shown in section 4.3.2 of the experiments.

$$\text{Percentage of road area} = \frac{\text{number of extracted pixel}}{\text{image width} \times \text{image height}} \times 100 \quad (4.3)$$

Where; image height and width were set by 800 and 600 respectively.

$$\text{Percentage of centroid } C_x = \frac{C_x}{\text{image width}} \times 100 \quad (4.4)$$

$$\text{case} = \begin{cases} LRA, \{(x, y) | 40 \leq x \leq 55 \text{ and } y \geq 25\} \\ \text{Others, otherwise} \end{cases} \quad (4.5)$$

Where; x = Percentage of centroid C_x and y = Percentage of road area

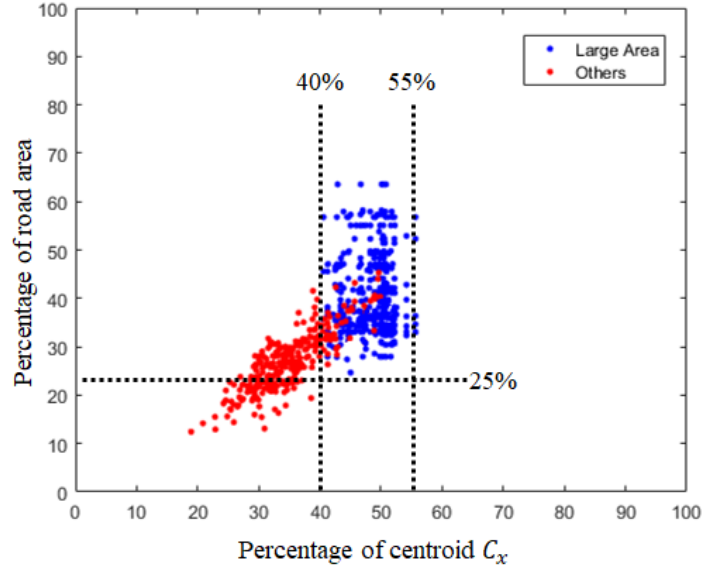


Figure 4.10 Comparison of plotting features between large road area viewpoints and other viewpoints.

4.2.3 Road Segmentation and Feature Extraction

Until now, the LRA is classified from other cases that consist of suitable and unsuitable viewpoints. Thus, the next process comprises the method for road segmentation and feature extraction.

4.2.3.1 Road Area Segmentation

Although the reference area of road can be found using the previous step, those segmented areas appear as a rough road area. Therefore, this section presents the technique that can be used to segment the road area clearly. For the road segmentation process, the two main steps comprise road edge detection using Hough Line Transformation (HLT) and line selection, followed by gray-scale and seed region growing segmentation.

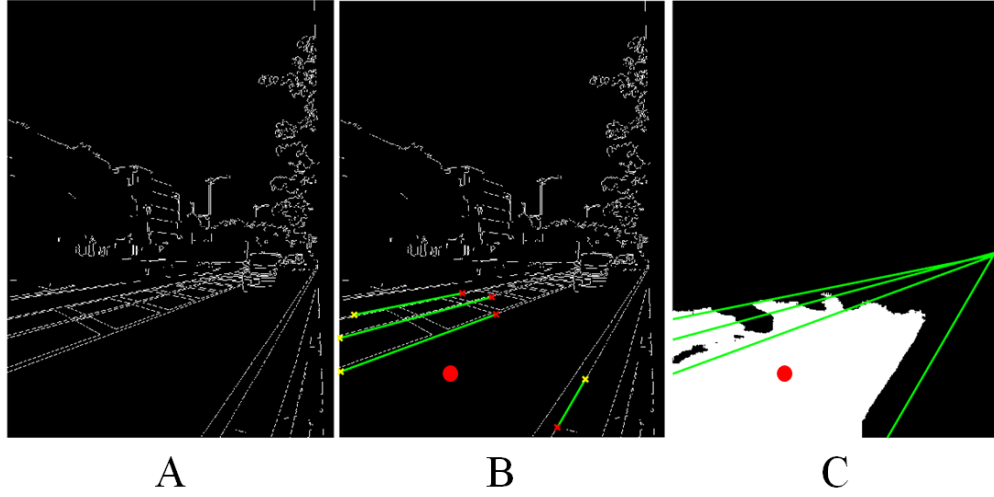


Figure 4.11 The step of road edge detection with Hough Line Transformation method.

A. Sobel edge detection, B. Line detection by HTL, C. Reference road area and feasible lines.

1. Road Edge Detection Using Hough Line Transformation and Line Selection

To border the road area, HLT is applied for detecting the straight-line markers on the road [45-47]. The HLT can transform each point in the (x, y) space to Hough space (θ, r) . Although, θ can be obtained from -90 to 90 degrees for HLT method, those are considered just 20 to 90 degrees because the possible marker lines on the image usually appear with those θ ranges for this proposed application. Figure 4.11A shows the image edge detection, which uses Sobel edge detection [48], which is a necessary step for the HLT transformation technique. The HLT technique is then used to detect the feasible lines, as shown in Figure 4.11B. In Figure 4.11C, many possible lines appear in the image with the reference road area (white pixel). Next step, two border lines of road marker will be selected by our proposed method.

According to previous step, the initial road (*reference road area*) has been detected and the centroid of the road pixel area is calculated as the red point in Figure 4.11C. The shortest distance between centroid and those lines need to be determined. First, it can be assumed that the (h, k) in Figure 4.12 is the centroid of road area. In addition, L_1 represents the straight-line that has been detected by HLT previously. The point (x, y) , which is the touching point

between L_1 and circle edge with perpendicular line of L_2 , which can be found using Eq. 4.6. Equation 4.6, m_1 and m_2 are the slopes for L_1 and L_2 , respectively. Further, the distance, (r), between the centroid and (x, y) coordinate points can be calculated using Eq. 4.7.

$$m_1 \times m_2 = -1 \quad (4.6)$$

$$r = \sqrt{(x - h)^2 + (y - k)^2} \quad (4.7)$$

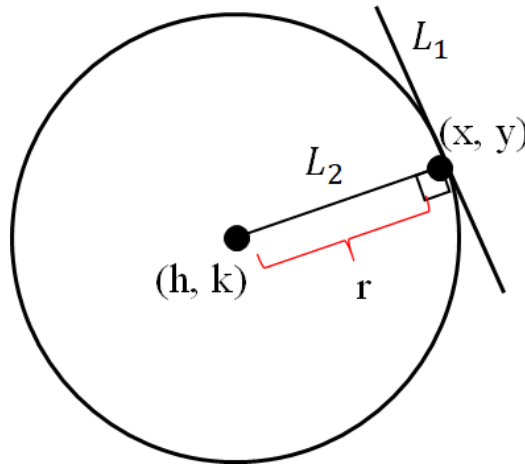


Figure 4.12 Line touching on the circle's edge.

Each line is calculated using the proposed method. The two lines are required, thus, the graph is separated into two areas comprising an upper and a lower centroid, as shown in Figure 4.13. For another case, when the road marker on the one side of the road is not clear because some obstacles or cars might be obscured, in Figure 4.14A. The algorithm can only detect the one side of road marker, and only case of upper centroid will be considered. Similar to two-line detected case, the shortest distance of upper centroid case will be selected for actual line on the road, as shown in Figure 4.15B.

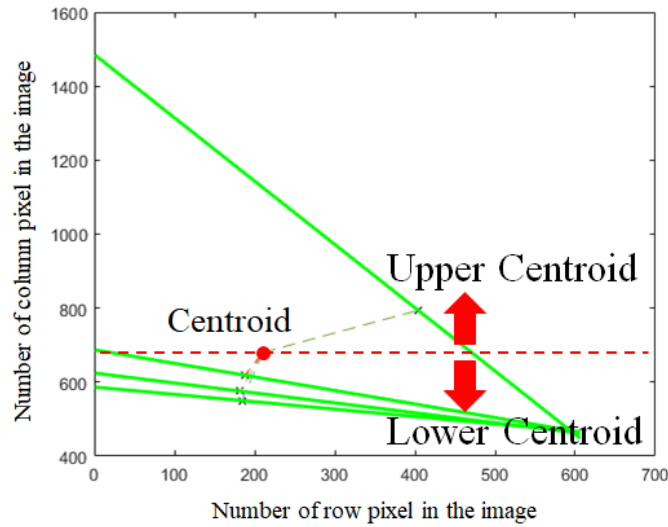


Figure 4.13 Upper and Lower centroid separation for two-line detection.

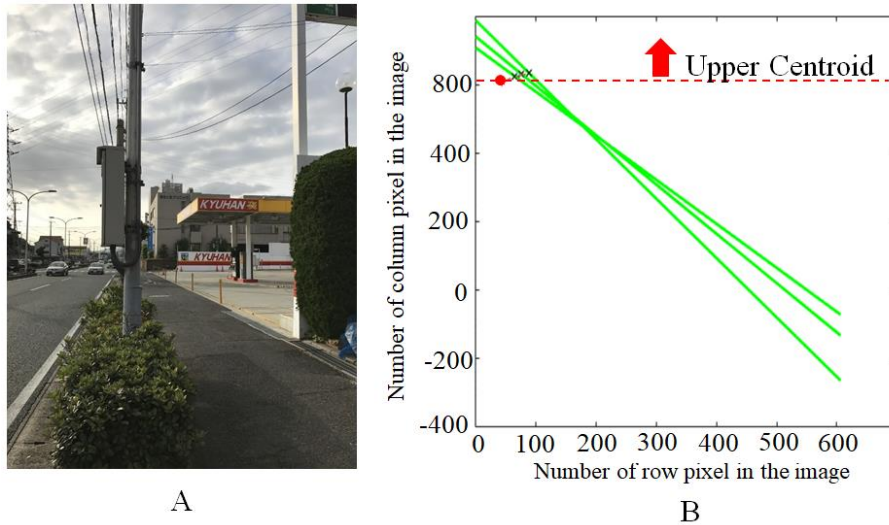


Figure 4.14 Case of one-line detection.

A. Example of unclear on road's marker, B. Only upper centroid detection for one-line detected case.

2. Gray-Scale and Seed Region Growing Segmentation

After defining the road area by detecting two lines, the next step is road segmentation. Gray-scale thresholding for outdoor segmentation is difficult to use, so this proposed method can be applied because the reference road area is roughly known. In Figure 4.15, an example

of gray-scale histogram of reference road area, previous process, is shown. Since, some reference road areas might contain some small noises such as a small part of a white marker on the road or some small non-road area; these small noises usually appear as the small peak, as shown in Figure 4.15A. Thus, this study sets the cut-off value at 90% differing with the highest peak for both sides, as shown as an example in Figure 4.15B. Figure 4.16 shows other examples for different histograms of gray-scale values that depend on illuminations of the reference road area are shown from the previous process. Subsequently, the remaining range of the histogram will be applied as the threshold value.

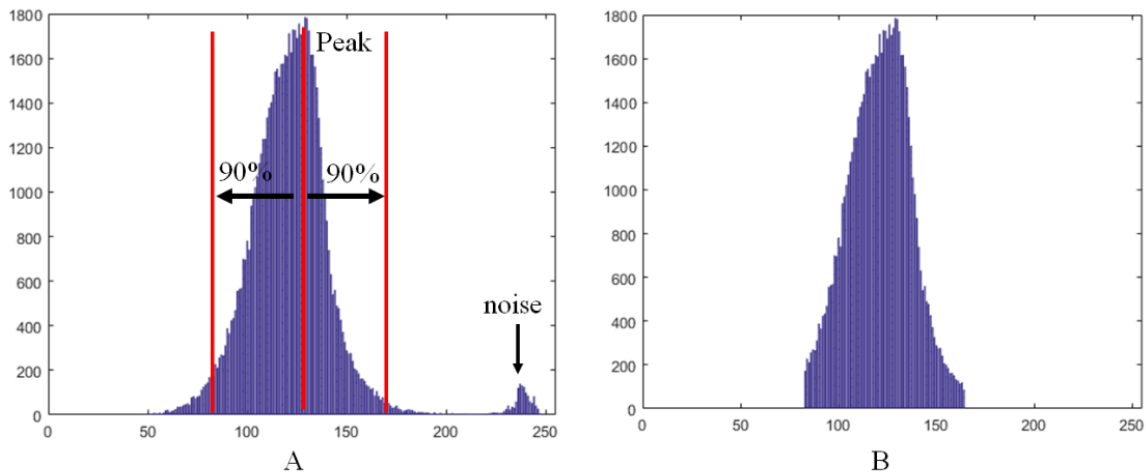


Figure 4.15 Setting cut-off value for noise elimination.

A. The original gray-scale histogram of reference road area, B. The gray-scale histogram after eliminating noise.

There are two cases of line detection, namely two-line detection and one-line detection. In the case of two-line detection, the gray-scale thresholding will be used for road segmentation within the enclosed area by selecting two lines, as shown in Figure 4.17A. Moreover, Figure 4.17B. shows the road segmentation with the threshold value, in case of two-line detection. Another case of one-line detection, in Figure 4.17C., is applied by the seed region growing technique [49-50] to segment the road area as shown in Figure 4.17D. The initial seed is the centroid of the reference road area. Each seed will extend to the four nearby directions with the threshold value mentioned in the previous proposed technique.

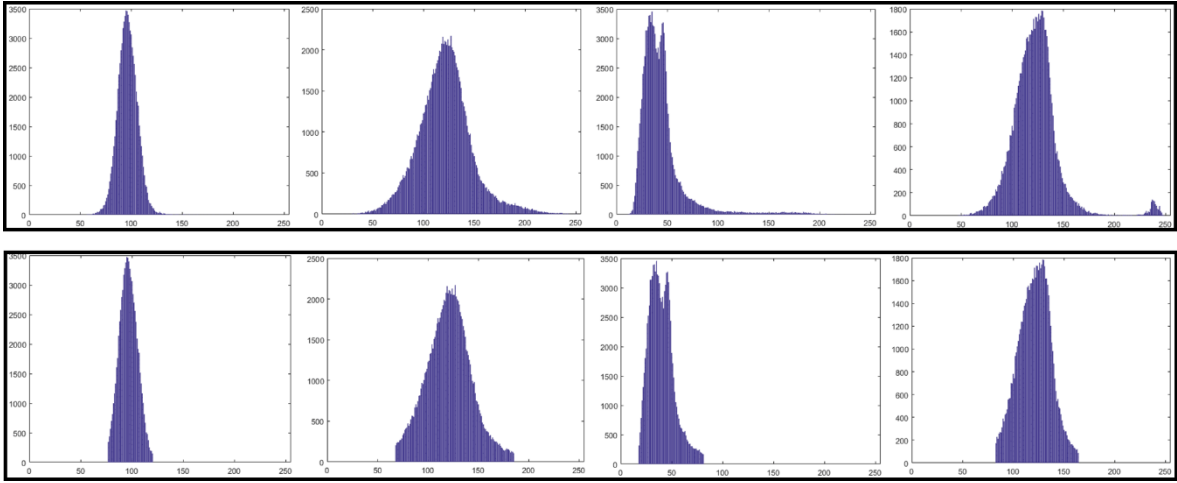


Figure 4.16 Example of gray-scale histogram with different illumination on the reference road area. The original gray-scale histograms (top row) and cut-off the histogram range by 90% differing with the highest peak for both side (lower row).

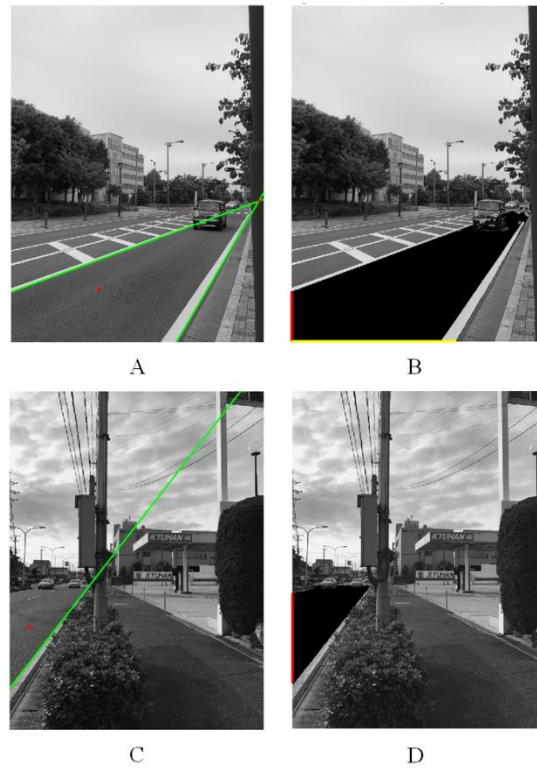


Figure 4.17 Gray-Scale and Seed Region Growing Segmentation.

A. Two-line detection case, B. Road segmentation using gray-scale technique, C. One-line detection case, D. Road segmentation using Seed region growing.

4.2.3.2 Feature Extraction

In order to classify the viewpoints for bus-waiting, this section will explain the necessary features based on the perspective view with the road area segmentation. This study proposes the necessary road features that consist of four features as shown in Figure 4.18. In Figure 4.18A and Figure 4.18B, three road features that comprise the vertical length of road (V), horizontal length of road (H), and the road areas can be calculated. Moreover, vanishing point (V_p), as shown in Figure 4.18A, is added for the case of two-line detection.



Figure 4.18 The necessary road-features for viewpoints classification.

A. Suitable viewpoint with features, B. Unsuitable viewpoint with features.

The definition of suitable and unsuitable viewpoints discussed, in chapter 3, are extracted and calculated in terms of the ratio of percentage of vertical length (V) and percentage of horizontal length (H), the percentage of vanishing point for the x coordinate, the percentage of vanishing point for the y coordinate, and the percentage of road area. Furthermore, Table 4.1 shows the list of features extracted with their significances of trend

of value between suitable and unsuitable viewpoints. The ratio of percentage of V and H is usually very small and large for suitable and unsuitable viewpoints, respectively. The percentage of V_p for x and y coordinates of suitable viewpoints are larger than the unsuitable viewpoints. In addition, the percentage of road area usually shows the small value for unsuitable viewpoints but a large value for suitable viewpoints. However, the V_p for x and y coordinates in case detection of one line will be set as zero value because it cannot find the vanishing point.

Table 4.1 List of features extracted with their significance.

Cases of lines detection	Features	Trend of value	
		Suitable viewpoints	Unsuitable viewpoints
1 and 2	% of V / % of H	Very small	Very large
2	% of V_p for x-coordinate	Large	Small
2	% of V_p for y-coordinate	Large	Small
1 and 2	% of road area	Large	Small

4.2.4 Viewpoints Classification based on Non-Congested Traffic

For the classification process, the ANN of feedforward backpropagation is applied, as similar as the Figure 4.7, which the input layer, the hidden layer and the output layer of ANN are designed by four, ten, and one nodes respectively. The four nodes of the input layer consist of the four feature values as shown in Table 4.1. Furthermore, the output layer is designed with one node that can return the value of one or zero for suitable viewpoints or unsuitable viewpoints, respectively. However, the performance of feature extraction and viewpoint classification with proposed method will be shown in section 4.3.4 of the experiments.

4.3 Experiments and Results

Since, this proposed application will be used for outdoor scenarios, roadside. There are various situations and illuminations depending on times and weathers between day and night. Therefore, the experiments will test the performance of the each important processes. Based on the proposed method in section 4.2, five experiments are tested that consist of performance

testing for selecting the k_n by using ANN classification, accuracy of pre-viewpoint classification using criteria setting, feature extraction performance, viewpoint classification accuracy, and the comparison results of proposed method.

4.3.1 Performance Testing for Selecting the k_n by using ANN Classification

As the purpose of this research is to use a smartphone to find the suitable viewpoints for an oncoming bus, various conditions of the road that depend on different illuminations from daytime until nighttime can appear on an image, and effect to the color tone of road. In order to ensure the performance of the proposed technique for selecting the group of k -means as mentioned in section 4.2.1.2, the road area is examined in various illumination conditions.

Eight conditions were tested, in Figure 4.19, namely sunny, cloudy, before sunset (and sunrise), during raining in daytime, after raining in daytime, nighttime, during raining in nighttime, and after raining in nighttime. Likewise, the non-road areas were added for this testing in order to show the performance when ignoring non-road areas.

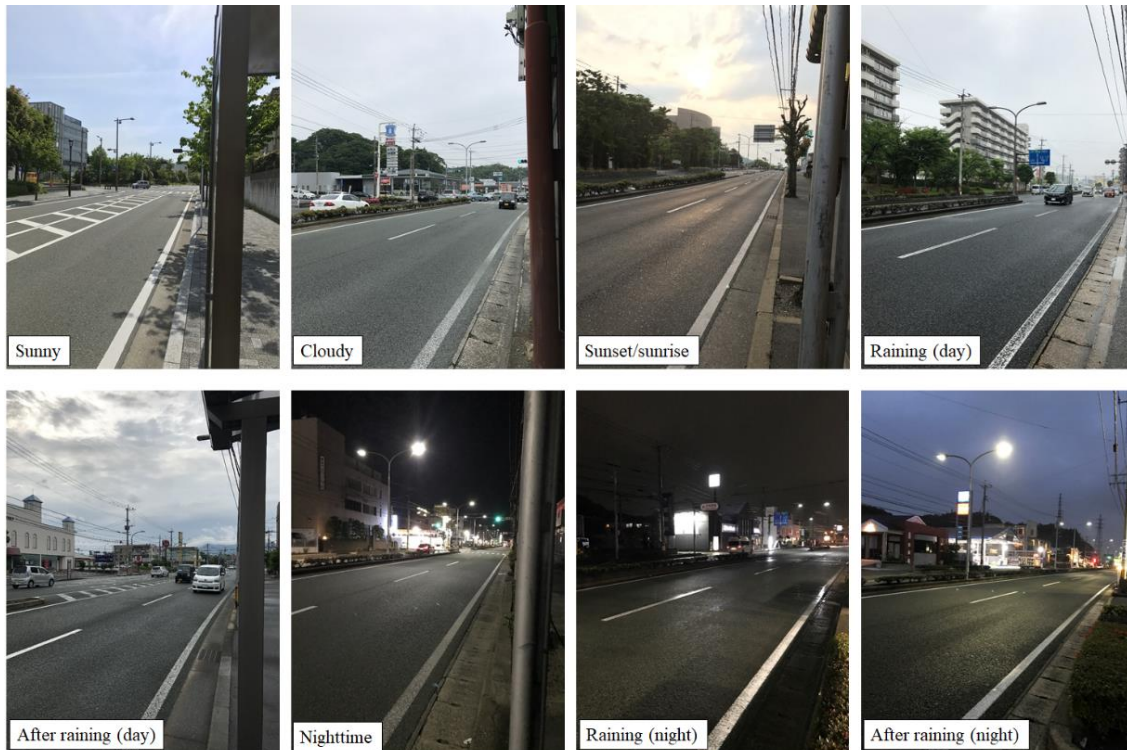


Figure 4.19 Eight different conditions between daytime and nighttime for experiment.

Table 4.2 Performance testing for selecting the group of k -means by using ANN classification.

Conditions	Accuracy (%)
Sunny	100
Cloudy	100
Before sunset / Sunrise	100
During raining in daytime	100
After raining in daytime	100
Nighttime	97
During raining in nighttime	97
After raining in nighttime	100
Non-road	96
Average of accuracy	98.89

For ANN training process, 400 different histograms for road and non-road areas, eight different conditions, were used. In addition, 1,800 different histograms were tested for eight conditions of road and non-road areas, which were tested by 100 images for each condition. In Table 4.2, the percentage of accuracy for nine different conditions, included the non-road area, is shown; the average accuracy is 98.89%; thus, this percentage shows high effectiveness of the proposed method to select the road area.

4.3.2 Accuracy of Pre-Viewpoint Classification using Criteria Setting

Based on the section 4.2.2, the excessive left panning viewpoint, that shows large area of road, will be pre-classified. The criteria value was set by Eq. 4.3, which applied the percentage of road area and the C_x coordinate of the vanishing point. This section shows the accuracy of pre-classification between large road area and other viewpoints.

Table 4.3 Accuracy of pre-viewpoint classification using criteria setting

Viewpoint types	Accuracy (%)
Large road area	98.00
Others	92.85

Eight hundred images for the large area of road viewpoints and other viewpoints were collected. Table 4.3 shows the experimental results of accuracy testing between large road area (LRA) viewpoints and other viewpoints. The accuracy percentages for LRAs and other viewpoints are 98 and 100, respectively. There are two percent errors for LRA viewpoints because a certain percentage of road area and the C_x coordinate of vanishing point are outside of the criteria setting. However, the accuracy is still high with the proposed criteria setting.

4.3.3 Feature Extraction Performance

Since, this proposed application will perform in various conditions, both of daytime and nighttime. This experiment will show the performance of feature extraction process as error percentage for eight different conditions.

As the previous section 4.2.3.2, five features were extracted such as vertical length of road, horizontal length of road, vanishing coordinate point ($V_p(x)$, $V_p(y)$) and area of road. Moreover, four hundred images were tested for all conditions, and five features were extracted for each image. Subsequently, the extracted feature values were compared with actual values that were measured by toolbox of AdobePhotoShop software. The error percentage was calculated as shown in Eq. 4.8.

$$\% \text{ error} = \left| \frac{\text{actual value} - \text{experiment value}}{\text{actual value}} \right| \times 100 \quad (4.8)$$

Table 4.4 shows error percentage and standard deviation (S.D.) for each condition and feature. All errors were lower than 7.00%, in both terms of conditions and features measurement. In term of conditions, the biggest error and S.D. are 6.36% and 10.67, which is when it is raining at nighttime. The roadside lights and vehicle headlights during rain can affect the unstable illumination of road area. Similar to sunny conditions, it shows the second highest error by 5.73% because this condition can appear various patterns of illumination from sunshine.

On the other hand, the condition after rain (daytime) shows the lowest error value by 0.96% and 1.14 of the S.D. value. This is because the road area after daytime rains is quite

homogeneous, owing to the road being wet without sunshine. Therefore, the performance of feature extraction is high in post rain conditions during daytime.

Table 4.4 Performance testing for feature extraction.

Conditions	% Error, S.D.					Average of % error, S.D. (Conditions)
	Vertical	Horizontal	Vp(x)	Vp(y)	Area	
Sunny	6.57, 10.17	11.40, 18.17	1.90, 4.02	1.05, 1.61	7.73, 10.44	5.73, 8.88
Cloudy	3.73, 8.48	10.56, 19.88	1.06, 1.02	0.92, 1.10	6.04, 8.38	4.46, 7.77
Before sunset/ Sunrise	2.81, 4.48	3.30, 5.69	1.35, 2.88	1.33, 1.83	3.33, 5.44	2.42, 4.06
During raining (daytime)	2.05, 4.75	5.94, 15.96	2.91, 6.23	1.91, 3.03	3.08, 4.30	3.17, 6.85
After raining (daytime)	0.81, 0.67	0.85, 0.97	0.96, 1.20	1.12, 2.01	1.06, 0.83	0.96, 1.14
Nighttime	1.40, 1.12	9.29, 17.23	4.08, 12.28	2.20, 4.03	5.78, 5.77	4.55, 8.09
During raining (nighttime)	5.25, 16.97	8.25, 10.13	4.78, 8.40	4.51, 5.97	8.99, 11.89	6.36, 10.67
After raining (nighttime)	1.28, 1.07	3.52, 7.43	0.86, 1.47	0.81, 0.78	2.66, 2.64	1.83, 2.68
Average of % error, S.D. (Features)	2.99, 5.96	6.64, 11.94	2.24, 4.68	1.73, 2.56	4.83, 6.21	

In terms of features, five extracted features will be applied for viewpoint classification. The highest error is 6.64% for horizontal length feature, but the feature of vanishing coordinate point ($V_p(y)$) shows the lowest error at 1.73% and vanishing coordinate point ($V_p(x)$) shows the second lowest error at 2.24%. This means that the performance of the vanishing coordinate point ($V_p(x)$, $V_p(y)$) extraction method is highly accurate compared with the other four features of extraction. Both error measurements for conditions terms and features terms show high performance, as seen in Table 4.4.

4.3.4 Viewpoint Classification Accuracy

As the main purpose of this chapter is to find the suitable viewpoints of bus-waiting

by the roadside without the obstacle consideration, the quantitative measurement is necessary. For this experiment section, all images were collected from the different bus stops around Wakamatsu-ku, Kitakyushu city, in Japan, with different condition for both of daytime and nighttime. The capturing photo area within 3 meters close to the road area, and within 3 meters before and next to bus stop area.

Table 4.5 Accuracy of viewpoint classification.

Conditions	Accuracy of viewpoint classification (%)		Average of % accuracy
	Suitable	Unsuitable	
Sunny	96	100	98
Cloudy	100	100	100
Before sunset/ Sunrise	100	100	100
During raining (daytime)	98	100	99
After raining (daytime)	100	100	100
Nighttime	96	100	98
During raining (nighttime)	96	100	98
After raining (nighttime)	99	100	99
Total average of accuracy	98.13	100.00	99.00
Standard deviation	1.89	0.00	0.93

The number of 400 images were provided for training the feedforward backpropagation of ANN. These training images include both suitable and unsuitable viewpoints that were considered as proposed techniques in chapter 3. After, all features were extracted by the proposed method in section 4.2.3.2 and data was sent to the ANN for training process. For the testing process, 800 images as unknown data were prepared for eight different conditions. In addition, each condition was half separated for suitable and unsuitable viewpoints. Then, ANN as previously mentioned, was used to classify all unknown data of images. The output of ANN showed the value (one or zero), and one or zero meant suitable or unsuitable viewpoints for bus-waiting, respectively. Subsequently, the output from the

ANN classification was compared with suitable definition, as explained in chapter 3.

According to Table 4.5, the quantitative results for final performance of viewpoint classification was presented. The experimental results show the highest accuracy of 100% in three conditions, namely cloudy, before sunset, and after raining in daytime. Nevertheless, the lowest accuracy was seen in three conditions of sunny, nighttime, and during rains in the nighttime by 98%. Further, the unsuitable viewpoints can be classified perfectly by 100% in all conditions. However, some mistakes in suitable viewpoint classification are seen whereby the total average accuracy is only 98.13%.

4.3.5 Comparison Results

Although, each step of our proposed have tested separately as above section, the performance of a whole proposed system is necessary to show the effectiveness. Moreover, the comparison results will be shown by this section.

4.3.5.1 Comparison with our three possible methods

Based on our proposed method, there are four main steps consisting of finding the reference of road area (step1), pre-classification (step2), road segmentation and feature extraction (step3), and viewpoint classification (step4) as shown in Figure 4.2. There are three possible combined methods for classifying the viewpoint based on our proposed.

Table 4.6 Comparison of three different combined methods based on our proposed algorithm.

Combined Methods	Accuracy (%)		
	Suitable viewpoint	Unsuitable viewpoint	Average
Step 1 + 2	79.77	27.33	53.55
Step 1 + 3 + 4	90.11	75.00	82.55
Step 1 + 2 + 3 + 4	98.13	99.00	98.56

First method is a combination between step1 and step2 of our proposed method. Second method of classification is integrated by step1, step2, and step4. For third method, a

whole step of our proposed is performed. Each combined method was used the 800 images of eight different conditions of times and weathers for testing the effectiveness. According to result in Table 4.6, the combined method of step1 and step2 showed 53.55% of average accuracy. Besides, the combined method without step2 shown the accuracy in average by 82.55%, and 98.56% for whole step of our proposed method.

4.3.5.2 Comparison with related research

To show the advantage of our proposed method for viewpoint classification in situation of bus-waiting. Although, there was no any research previously concerning the totally same as our proposed, the comparison result between our proposed and related research is necessary. A research of Hangrong P. *et al.* [19] concerned the similar situation and image viewpoint of our proposed for bus-waiting, roadside, between the possible and impossible to detect the oncoming bus, using the smartphone camera.

Table 4.7 Comparison result between our proposed method and related research.

Methods	Accuracy (%)		
	Suitable Viewpoint	Unsuitable Viewpoint	Average
Hangrong P. <i>et al.</i> [19]	81.48	80.19	80.83
Our proposed method	98.13	99.00	98.56

According to Table 4.7, the comparison result was shown. The average of accuracy for our proposed method was higher than the related research, which were shown by 80.83% and 98.56% for Hangrong P, *et al.* and our proposed method. In addition, the accuracies of our method were higher than the previous research for both of suitable and unsuitable viewpoint classification. Obviously, our proposed method has the effectiveness over than the previous research that shows the very high percentage of accuracy.

4.4 Discussion

This chapter proposed the novelty of viewpoint classification for non-congested traffic

without obstacles consideration. The advantage of this proposed viewpoint classification is to complement existing research that solely focused on bus detection and bus number recognition. Moreover, this proposed method applies the combined simple techniques of image processing. The proposed application will apply to smartphone with real-time process, therefore, the simple technique with light load calculation is feasible and suitable. Although, this research did not measure the exactly processing time, this proposed method take just a few second on MATLAB2017b with Core i7 of processor and 8.00GB for PC RAM (64-bit).

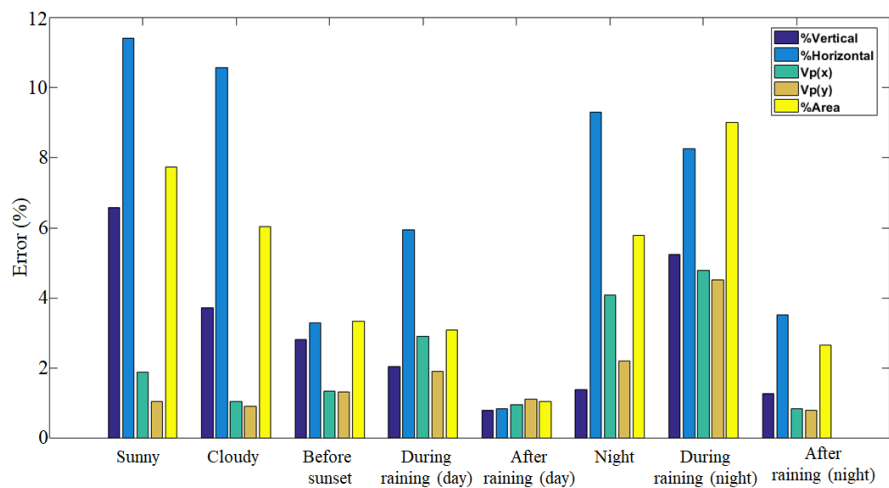


Figure 4.20 Graphical representation of error percentage for feature extraction.

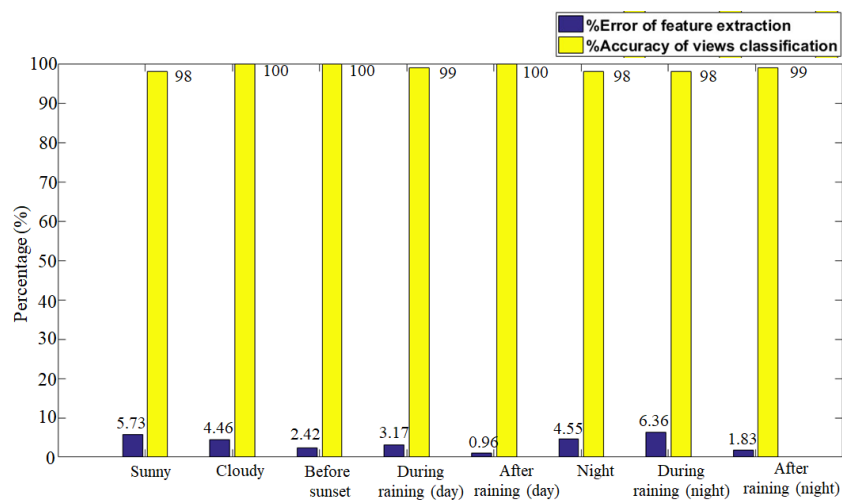


Figure 4.21 Compared graphical between viewpoint classification and error of feature extraction.

In this section, the differences between error results of feature extraction process, in Table 4.4, and accuracy of viewpoint classification, in Table 4.5, will be described as shown in Figure 4.20 and Figure 4.21, respectively. The error results for each condition from Table 4.4 are represented in Figure 4.20. The condition after raining (day) has the lowest error compared to others. Although an error in the feature extraction process appears in the graph, as shown in Figure 4.20, the accuracy of viewpoint classification is still high in all conditions, as shown in Figure 4.21.

As the lowest average of error value (0.96%) corresponds to the condition after rain in daytime, the highest accuracy for this condition is seen at 100%. In contrast, the lowest accuracy of viewpoint classification can be seen in three conditions of sunny, nighttime, and during rain in the night at 98%, which is the biggest error of feature extraction corresponding to rain in the nighttime at 6.36%. Furthermore, 5.73% and 4.55% are the second and third highest errors for sunny skies and raining at night conditions, respectively.

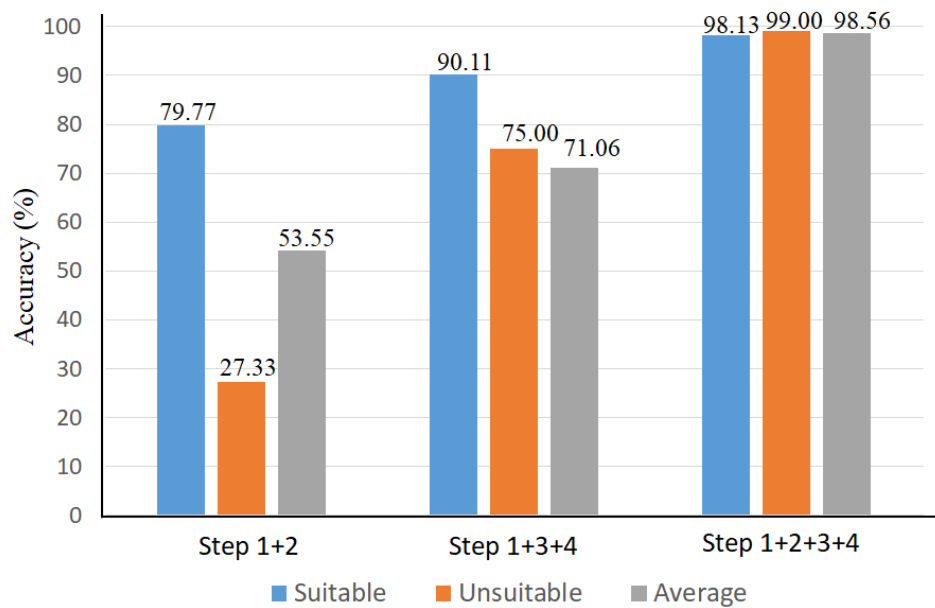


Figure 4.22 Comparison graph of three different methods of our proposed.

In section 4.3.5.1, the comparison of three different methods based on our proposed was shown. Figure 4.22 shown the comparison graph of different combined methods. For the

combination of step1 and step2, it shown the lowest accuracy, both of suitable and unsuitable viewpoint, compared to other combined method. Moreover, when the combined method without step2 was performed, the accuracy become higher than the case of step1 and step2 combination for both of suitable and unsuitable viewpoint classification. Obviously, the highest effectiveness of system was shown, when whole step was used. Therefore, this experiment confirm that all step of our proposed is crucial for viewpoint classification.

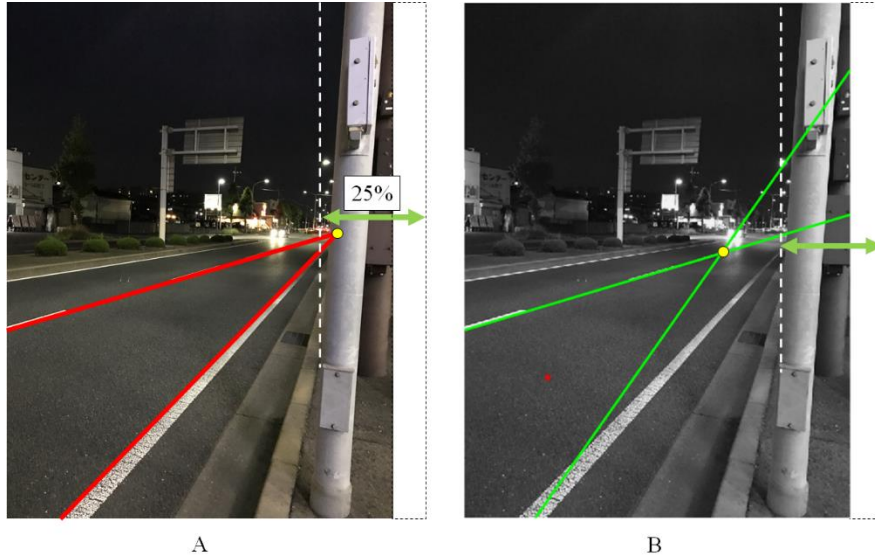


Figure 4.23 Example of error case for line detection.

A. Suitable viewpoints image for bus-waiting as defined in chapter 3, B. Example of mistake for line selection.

Although, these results show high performance, there are some mistakes in the viewpoint classification process. For suitable viewpoint classification, especially, wrong results can come about when comparing the proposed definition of suitable viewpoints for bus-waiting described in chapter 3. For example, in Figure 4.23, an example error case of viewpoint classification is shown. According to definition, the suitable viewpoints for bus-waiting are shown in Figure 4.23A. However, this classification of viewpoints is inaccurate because of wrong line selection, as shown in Figure 4.23B. The process of HLT can be detected, so many lines are shown in Figure 4.24, but can be mistaken for some cases of the line selected process.

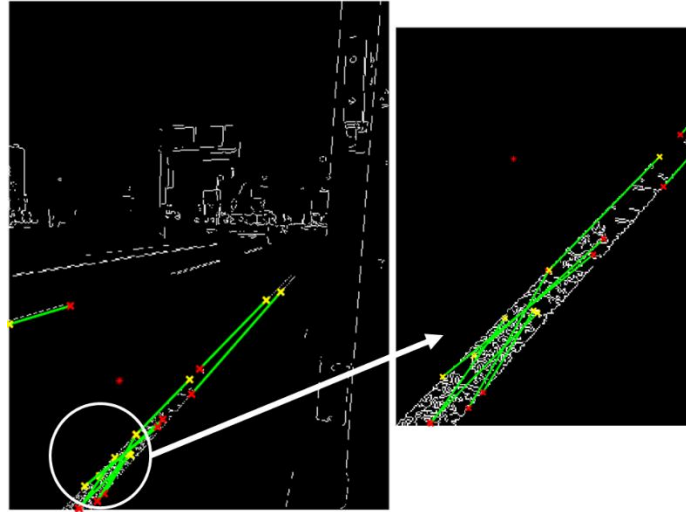


Figure 4.24 Many detected lines of HLT process.

4.5 Summary

This chapter proposed the novel application of viewpoint classification in cases of non-congested traffic conditions related to bus-waiting, an idea that can be of great help to blind individuals in Thailand.

The road area is detected for extracting the necessary features using a combination of methods. This technique is robust across various illuminations. There are four main processes that constitute the proposed method, namely finding the reference of road area, pre-viewpoint classification, road segmentation and feature extraction, and viewpoint classification.

Based on this method, real scenarios of roadside viewpoints for bus-waiting with eight different conditions in daytime and nighttime were tested. The experimental results show the high accuracy of viewpoint classification by 98.56%.

This page intentionally left blank

5. Classification of Viewpoints while Waiting for the Bus in Situation of Congested Traffic

The situation of congested traffic usually appears on the road, especially the big city as Bangkok, Thailand. Therefore, this situation will be explained in this chapter. The chapter includes five sections. First section explains the objective of this study. Second, the suitable and unsuitable viewpoint estimation in case of congested traffic will be explained based on suitable viewpoint definition as mentioned in chapter 3. Next, the proposed method for classification of viewpoint in situation of congested traffic without obstacle consideration, which the car distribution will be considered, is shown. Then, the experiments and results, in section 5.4, will be explained with the optimized results of features selection and classifiers selection. Finally, the result and the proposed methods will be discussed in the last section of this chapter.



Figure 5.1 Difference between non- and congested traffic viewpoint.

5.1 Objective of Classification of Viewpoints in case of Congested Traffic

The main objective of this study is to classify the viewpoint in case congested traffic situation while the blind waiting for the bus for daytime and nighttime. The congested traffic

is one situation that usually appears in the big city. Especially, the rush hour for going to work or school in the morning, and going back to their home after working. Therefore, this study will focus and discuss in this situation for helping blind people.

Although, chapter 4 has proposed the classification for non-congested traffic based on the road-feature extraction, the congested traffic condition cannot apply as the road-feature consideration. Since, many cars as shown in Figure 5.1 will obscure the road area and road line-marker sometimes. This chapter will propose another method based on the car distribution consideration for classifying viewpoint.

5.2 Viewpoint Estimation in case of Congested Traffic based on Suitable Viewpoint Definition

As the suitable viewpoint definition in chapter 3, the definition was described by the non-congested traffic image in Figure 3.14 and Figure 3.15. However, we can define and estimate the suitable and unsuitable viewpoint of congested traffic situation as same as the non-congested traffic situation as shown in Figure 5.2. Manually, the vanishing point (V_p) was defined by drawing the two convergent lines. Then, the V_p was compared to the suitable viewpoint definition within the 25% range (yellow area), as explained in chapter 3.

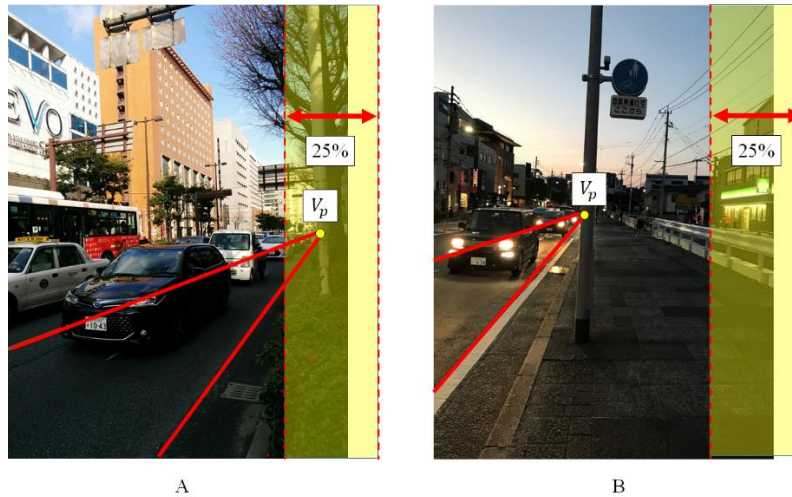


Figure 5.2 Estimation of suitable and unsuitable viewpoint in case of congested traffic based on suitable definition.

A. Suitable viewpoint, B. Unsuitable viewpoint.

5.3 Proposed Method for Classification of Viewpoint based on Congested Traffic

This chapter aims to classify the viewpoints of blind individuals who waiting for buses by the roadside, especially in the case of congested traffic. Moreover, in order to classify the suitable panning of the camera, this study assumes that all images were taken by holding smartphones vertically. According to the suitable-viewpoint definition in chapter 3, the vanishing point (V_p) comprises two convergent lines on the road; however, it is difficult to find these lines in congested traffic automatically because there are many cars on the road obscuring both of road line marker and road area. Further, the proposed application have to perform in daytime and nighttime as shown in Figure 5.3.



Figure 5.3 Congested traffic viewpoints under daytime and nighttime.

Consequently, our method proposes a technique that finds the car distribution in an image related to definition. The four main steps of the proposed method consist of car detection using the YOLO technique, data normalization, feature extraction, and viewpoint classification, as shown in Figure 5.4.

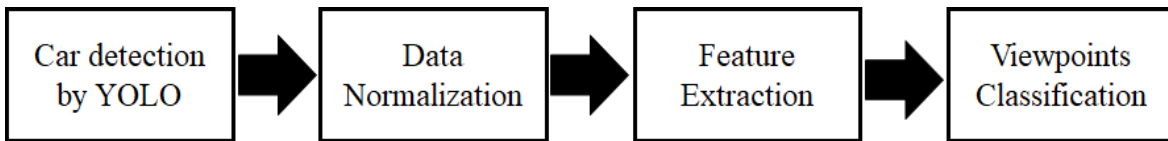


Figure 5.4 Four main processes of proposed method for viewpoint classification in case of congested traffic.

some farther cars were unnecessary, 150 px was set for the distance criterion between the x, y coordinate points at the top-left and bottom-right (original image size is 1,944×2,592). The Table 5.1 shows the parameter setting for YOLO that applied to car detection.

Table 5.1 Parameter setting of YOLOv2 for car detection.

Parameters	Criteria setting
Object labels	Car, truck, bus
Confidence value	0.38
Boundary size (Min.)	150 px

After setting all parameters, YOLO was applied for prepared images showing various real situations under congested-traffic conditions. Although the cars detected using YOLO provide the square boundary, as shown in Figure 5.6, the proposed method solely uses the x, y coordinates of the center boundary as a blue point. All center points were provided for the feature-extraction process. Moreover, the section 5.4.1 shows the car detection performance using this proposed method and criteria setting.



Figure 5.6 Outcomes of car detection using the YOLOv2 technique (these images are related to the graphs of Figure 5.7).

5.3.2 Data Normalization

In order to provide the data for feature extraction, data normalization was necessary

because of different sizes of original-input images. Each center point (x_n, y_n) of a detected car was normalized as a percentage. Eq. 5.1 and Eq. 5.2 show the normalized calculation for x'_n, y'_n , and Figure 5.7 shows the outcome of this step.

$$x'_n = \frac{x_n}{\text{number column of original image}} \times 100 \quad (5.1)$$

$$y'_n = \frac{y_n}{\text{number row of original image}} \times 100 \quad (5.2)$$

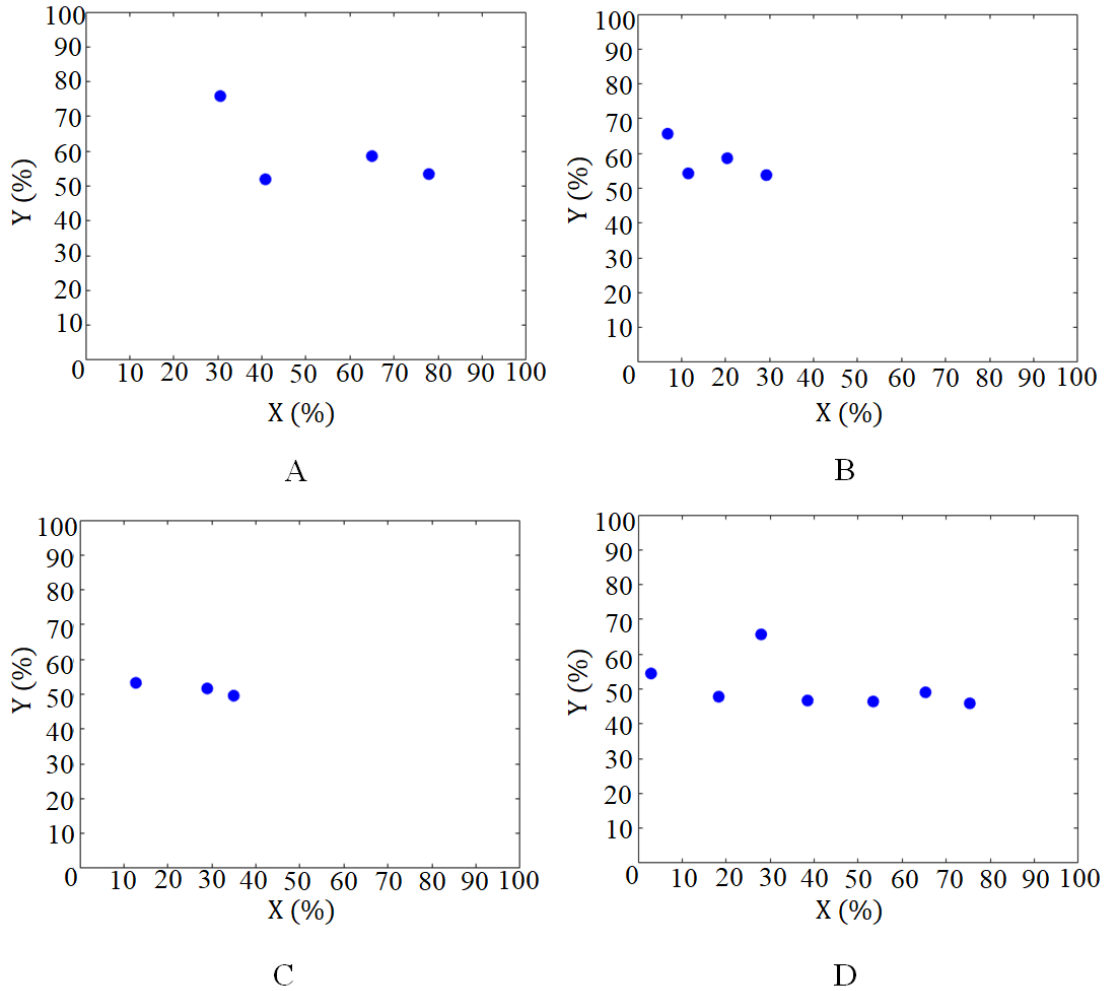


Figure 5.7 Data normalization.

(A) and (D) example of data normalization for suitable viewpoints, (B) and (C) example of data normalization for unsuitable viewpoints.

5.3.3 Car Distribution-Feature Extraction

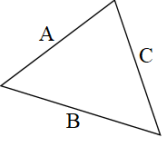
The previous study [53] extracted thirteen features, but some possibly useful ones may have been neglected. In addition, the car distribution in each image appeared randomly depending on different congested-traffic situations, as shown in Figure 5.7, and there are a few data points that can be calculated for the feature-extraction process. Therefore, this section will explain and present the possible features based on car distribution in the image.

Each image was calculated for providing the feature vector. There were nineteen possible features extracting from datapoints after normalization process, which were extracted as statistics calculation, coefficient of linear regression, and geometric value. Table 5.2 presents the list of all features from normalized datapoints in the images.

Table 5.2 List of features and their descriptions

Features	Description
Number of cars	$n X = \{x_1, \dots, x_n\}$ and $Y = \{y_1, \dots, y_n\}$
Maximum value of x and y	$x_{max} = \max(X), y_{max} = \max(Y)$
Minimum value of x and y	$x_{min} = \min(X), y_{min} = \min(Y)$
Mean of x and y	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}, \bar{y} = \frac{\sum_{i=1}^n y_i}{n}$
Median of x and y	$med_x = X_{(\frac{n+1}{2})}, med_y = Y_{(\frac{n+1}{2})}$
Standard deviation of x and y	$sd_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2},$ $sd_y = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2}$

Table 5.2 List of features and their descriptions (cont.)

Features	Description
Range of x and y	$R_x = x_{max} - x_{min}$ $R_y = y_{max} - y_{min}$
Center of range for x and y	$CenR_x = \frac{R_x}{2}, CenR_y = \frac{R_y}{2}$
Slope (m), y-interception (c) of linear regression	$m = \frac{n \sum_i^n x_i y_i - (\sum_i^n x_i \sum_i^n y_i)}{n \sum_i^n x_i^2 - (\sum_i^n x_i)^2}$ $c = \bar{y} - m\bar{x}$
R^2 of linear regression	$R^2 = 1 - \frac{\sum_i^n (y_i - \hat{y}_i)^2}{\sum_i^n (y_i - \bar{y})^2}$ <p>\hat{y}_i represents the y value of linear regression.</p>
The biggest of triangle area 	$Area = \sqrt{S(S - A)(S - B)(S - C)}$ $S = \frac{A + B + C}{2}$

Since there are different numbers of datapoints depending on the number of detected cars in the image, some features could not be calculated. For one and two detected cars, which the datapoints will appear 1 and 2, the biggest of triangle area feature will be set as zero value because the data point is not enough for calculation. However, when the datapoints are over than three points, the triangle can be calculated. Then each triangle area will be compared, and the biggest triangle will be selected for feature. In addition, when only one car was detected, the standard deviations, ranges, and centers of these ranges for x and y, as well as the slope, y-intersection, and R^2 of linear regression were set to zero. However, these nineteen features were just all feasible-feature values from the car distribution. In section 5.4.2 shows

the optimization for features selection and its performance.

5.3.4 Viewpoints Classification based on Congested Traffic

Since, there were no previous related researches proposing the viewpoint classification of bus-waiting in case of congested traffic situation. The classifiers are an important for performance of proposed method. This study used the supervised machine learning for the classification process. Five general different types of supervised machine learning were selected as following; 1) simple-decision tress, 2) Random forest, 3) Naïve Bayes, 4) Multi-layer perceptron, and 5) Support-vector machine. Moreover, the initial parameters for each classifier, using WEKA 3.8 software [54], were set as Table 5.3 to Table 5.7.

Table 5.3 Initial parameter setting for simple-decision tress

Parameter names	Parameter values
batchSize	100
binarySplits	False
collapseTree	True
confidenceFactor	0.25
minNumObj	2
numDecimalPlaces	2
numFolds	3
seed	1
subtreeRaising	True
unpruned	False
useMDLcorrection	True

Table 5.4 Initial parameter setting for Random forest

Parameter names	Parameter values
bagSizePercent	100
batchSize	100
breakTiesRandomly	False
calcOutofBag	False
maxDepth	0
numDecimalPlaces	2
numExecutionSlots	1
numFeatures	0
numIterations	100
seed	1

Table 5.5 Initial parameter setting for Naïve Bayes

Parameter names	Parameter values
batchSize	100
numDecimalPlaces	2
useKernelEstimator	Flase
useSupervisedDicretization	Flase

Table 5.6 Initial parameter setting for multi-layer perceptron

Parameter names	Parameter values
batchSize	100
hiddenLayers	9 nodes
learningRate	0.3
momentum	0.2
nominaltoBinaryFilter	True
normalizeAttributes	True
normalizeNumericClass	True
numDecimalPlaces	2
trainingTime	500
validationSetSize	0
validationThreshold	20

Table 5.7 Initial parameter setting for Support-vector machine

Parameter names	Parameter values
batchSize	100
buildCalibrationModels	False
c	1.0
calibrator	Logistic regression
epsilon	1.0E-12
filterType	Normalize training data
kernel	Polynomial Kernel (exponent 1.0)
numDecimalPlaces	2
numFolds	-1
randomSeed	1
toleranceParameter	0.001

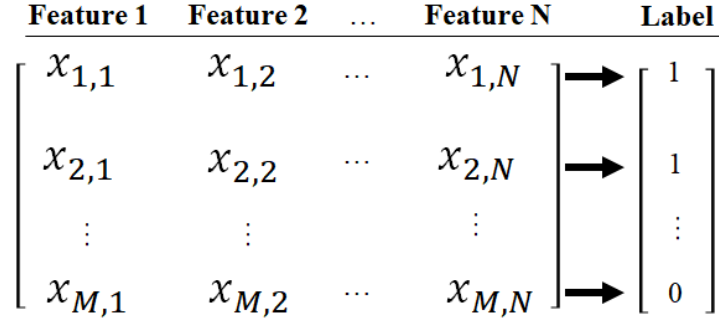


Figure 5.8 Feature-matrix arrangement and its labels.

Generally, there are two main steps of supervised learning, namely training and testing. All feature vector was separated into two classes as suitable and unsuitable viewpoints, following the definition in chapter 3. In Figure 5.8, a feature matrix shows N selected features. In addition, M represents the number of datapoints, which was four hundred in this study. Each datapoint was labeled as 1 or 0 for suitable and unsuitable viewpoints, respectively.

In order to find the best classification result between the classifier and the feature-selected method, each classifier was tested with different numbers of features, as show in section 5.4.2.

5.4 Experiments and Results

There were two experiments for this study, which consisted of car detection and viewpoint classification. In order to evaluate the performance of the proposed method, four hundred images under both day and night conditions, in real congested-traffic situations, were taken. The original images were in an RGB-color format with different sizes because we used three smartphones to collect them.

5.4.1 Car-detection Performance

The car-detection process was a crucial step of the proposed system because this study considered the car distribution in image for viewpoint classification. YOLOv2 method was used for car detection with the parameter setting mentioned in Table 5.1. Herein, the outcome

performance was tested by 400 images for day and night.

Table 5.8 shows the results of the detection process with 79.90% accuracy, as was compared to the actual number of cars counted by humans. However, the accuracy of the nighttime situation was lower than during the daytime (76.74% and 83.72%, respectively), because of visibility can be difficult at nighttime due to darkness and the headlights of cars.

Table 5.8 Car-detection performance using YOLO and the proposed parameter setting.

Conditions	Number of cars		Accuracy (%)
	Actual	Detected	
Day time	860	720	83.72
Night time	1,106	851	76.74
Day and night times	1,966	1,571	79.90

5.4.2 Optimization for Features and Classifiers Selection

According to section 5.3.3, possible nineteen features have been provided for the proposed method. Moreover, five different classifiers of supervised learning were selected for classification process, as mentioned in section 5.3.4. This experiment will find the highest accuracy with each matching feature and classifier. This experiment applied the WEKA 3.8 software [54] developed by the University of Waikato, New Zealand, because it offers useful feature-selection and classification tools.

For the experiment, 10-fold cross-validation was applied for testing the accuracy of classification process. The cross-validation technique is widely used for data classification, for which the whole dataset is divided into training and evaluation datasets repeatedly. Moreover, 10-fold cross validation means the dataset is partitioned into 10 equal subsamples. Then, one part of partitioned dataset was evaluated while others were used for data training. The process was repeated ten times, and the average of ten results was used as the final accuracy of the classification step.

Table 5.9 Comparison of accuracy with different feature selections and classifiers.

Feature selection methods	Number of selected features	Accuracy (%)				
		J48	Random forest	Naïve Bayes	Multi-Layer Perceptron	SMO
No selection	19	81.00	86.00	78.00	85.00	78.75
CfsSubsetEval	6	76.75	85.75	82.50	83.25	75.50
Correlation	13	78.25	85.25	77.25	81.50	77.25
InfoGain	17	81.50	86.00	78.50	83.00	77.00
OneRAttribute	10	80.00	84.75	77.00	80.50	76.25
PCA	6	75.00	81.25	76.75	85.50	79.25

Moreover, the experimental results show the six different methods of features selection and five general classifiers of supervised machine learning.

Table 5.9 shows the feature-selection methods comprising no selection, CfsSubsetEval, correlation, information gain (InfoGain), OneRAttribute, and principal components analysis (PCA). The following five classifiers were used: simple decision tree (J48), random forest, naïve Bayes, multi-layer perceptron, and the support-vector machine for sequential minimal optimization (SMO). Each classifier was tested by different feature-selection methods, with the highest outcome shown by the random forest classifier with no selection and InfoGain. Although, these two cases both showed 86.00% accuracy, the number of selected features of InfoGain was smaller than any other (namely, 17).

Therefore, to attain the best performance for our proposed method, this study selected a random forest classifier and the seventeen features shown in Table 5.2, excluding the mean of y and the y-intersection features.

5.4.3 Performance Measurement for Viewpoint Classification

After the classifier and features were selected, the confusion matrix of viewpoint classification is shown in Table 5.10. Since there were two classes of label, namely suitable and unsuitable viewpoints, the confusion matrix generally provides true-positive (TP), false-positive (FP), false-negative (FN), and true-negative (TN) results. Based on this information, the precision and F – measure can be calculated by Eqs. 5.3, 5.4, and 5.5 respectively:

$$\text{recall} = \frac{TP}{\text{Actual suitable viewpoint}} = 0.89 \quad (5.3)$$

$$\text{precision} = \frac{TP}{\text{Predicted suitable viewpoint}} = 0.84 \quad (5.4)$$

$$\text{F – measure} = \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}} = 0.86 \quad (5.5)$$

All outcomes were quite high by 0.89, 0.84 and 0.86 for recall precision and F – measure respectively. According to these results, it was found that the proposed method performed better than that from prior research [53].

Table 5.10 Confusion matrix for viewpoint classification.

		Actual		
		Suitable viewpoints	Unsuitable viewpoints	
Predicted	Suitable viewpoints	TP = 178	FP = 34	212
	Unsuitable viewpoints	FN = 22	TN = 166	188
		200	200	400 (Total)

5.4.4 Comparison Result

To show the benefit of our proposed method for viewpoint classification of bus-waiting, especially, the case of congested traffic. Although, there was no many researches previously concerning the perfectly same as our proposed, the comparison result between our proposed and related research is necessary. A research of Hangrong P. *et al.* [19] concerned the similar situation and image viewpoint of our proposed for bus-waiting, roadside, between the possible and impossible to detect the oncoming bus, using the smartphone camera. This comparison for our proposed method used four hundreds images between suitable and unsuitable viewpoints.

Table 5.11 Comparison result between our proposed method (congested traffic) and related research.

methods	Accuracy (%)		
	Suitable Viewpoint	Unsuitable Viewpoint	Average
Hangrong P. <i>et al.</i> [19]	81.48	80.19	80.83
Our proposed method	86.45	83.55	86.00

According to Table 5.11, the comparison result was shown. The average of accuracy for our proposed method was higher than the related research, which were shown by 80.83% and 86.00% for Hangrong P *et al.* and our proposed method. In addition, the accuracies of our method were higher than the previous research for both of suitable and unsuitable viewpoint classification. Obviously, our proposed method has the effectiveness over than the previous research with high percentage.

5.5 Discussion

Based on the results in section 5.4, the performances of car detection and viewpoint classification were tested under daytime and nighttime illuminations. The average accuracy of car detection was 79.90% in Table 5.8, and was lower at nighttime than at daytime due to darkness and the shine from car headlights. Furthermore, in order to remove some cars

located on other lanes of roads, the square-size parameter was set to 150 px. Then, the number of detected cars was smaller than the number of actual cars seen in Table 5.8. However, detection accuracy may be improved by adjusting some parameters of the YOLO algorithm, such as confidence value.

For viewpoint classification, the random forest classifier showed the highest performance with 86% accuracy. Moreover, seventeen features were chosen, as shown in Table 5.2, without the y-mean or y-intersection of the linear regression. In addition, the recall and precision were 0.89 and 0.84 respectively, which were both quite high. The F – measure shows the relationship between recall and precision, and thus had a high performance of 0.86.

Although, the calculation time did not measure the definitely value, this proposed method applied fast calculated algorithm of YOLOv2 (40 frames per second) and simply feature calculation. Therefore, it appropriate to use for the real-time application in future work.

5.6 Summary

This chapter proposed a novel viewpoint-classification application for assisting blind individuals using computer-vision techniques while waiting for buses. Especially, classification under congested traffic was considered. The proposed algorithm had four main steps, namely car detection, data normalization, feature extraction, and viewpoint classification. YOLOv2 was used for car detection because it can be implemented fast in real-time. Moreover, all centers of detected cars were normalized, and nineteen features were extracted. In order to classify the viewpoints, this study applied supervised learning. Based on the experimental results, the car-detection performance showed 79.90% accuracy. The classification accuracy was compared between different feature-selection methods and classifiers. The results showed seventeen features and the random-forest classifier provided the highest accuracy by 86.00%. Additionally, the recall, precision, and F – Measure were shown as 0.89, 0.84, and 0.86 respectively. The proposed method is considered to be feasible for real-time implementation with high performance in future work.

This page intentionally left blank

6. Obstacle Detection along the Road

This chapter considers the obstacle detection along the road, in case of non-congested traffic situation. The chapter includes five main sections. First, the objective and importance of this study will be explained. Second, the properties of obstacle along the road, which might obscure the suitable viewpoint for detecting the route bus number, will be discussed. Third, the method for the static obstacle detection along the road is proposed that consists two main steps of obstacle's position finding and obstacle's height estimation. Fourth, the experiments and results of proposed method, both of two main steps, is shown. Finally, the results and the proposed methods will be discussed in the last section of this chapter.

6.1 Objective of the Obstacle Detection along the Road

Non-congested traffic situation without obstacle appearance have been considered, in chapter 4. However, obstacles may appear in the images when the image is captured. This chapter defines the meaning of obstacle along the road such as electricity poles, poles, tree, any static objects, and human. In real-world conditions, obstacles have various shapes and sizes as shown in Figure 6.1. Moreover, some obstacles as the tall or big object that could be captured in images, oncoming buses or their route numbers may not be detected. To achieve the final goal for our assistive application of suitable viewpoint classification, an obstacle consideration process is necessary.



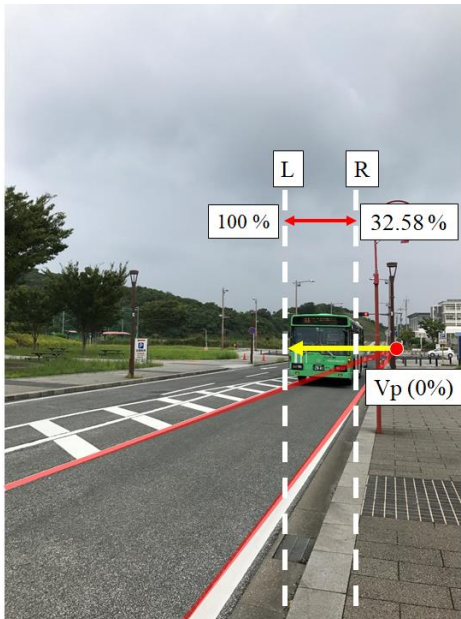
Figure 6.1 Obstacles along the road.

6.2 Properties of Obstacles

Although many obstacles can appear in an image, some of the obstacles can obscure the route number of the oncoming bus, which related to the definition of a suitable viewpoint, as mentioned in chapter 3. There are two main properties of obstacle relating to the suitable viewpoint definition. First property is a range of interest in the image that will obscure the oncoming bus. Second, the height of obstacle will be estimated for our system.

6.2.1 Range of Interest

For the first property, the system was designed with a range of interest (R-L) from 32.5% to 100% of the range, as shown in Figure 6.2, which begins at 0%, corresponding to the vanishing point of the perspective image.



Range of Interest calculation

1. The system defines the Vp and L line as the starting point (0%) and ending of range of interest (100%).
2. Then, the distance between Vp and R equals 32.58%.
3. The Range of Interest is within 32.58% to 100%.

Figure 6.2 Range of Interest for obstacles consideration.

6.2.2 Height of Obstacle

For the property of height for obstacle, the principle of perspective view [55] will be used. The principle of perspective view mentioned that the horizon line represents the eye level or camera position of users. As the Figure 6.3, the horizon line is shown, which the line

will always be on the same level with the vanishing point (V_p).

Therefore, the property for height of obstacle is defined by horizon line for our proposed system. Herein, if the obstacle is higher than the horizon line, it will obscure the suitable viewpoint. On the other hand, if the obstacle is lower than the horizon line, that obstacle will not effect to route bus number detection as shown in Figure 6.4.

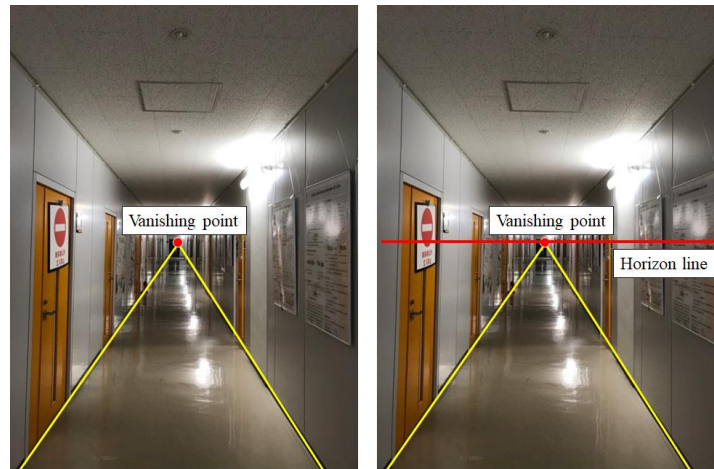


Figure 6.3 Horizon line and vanishing point of perspective image.

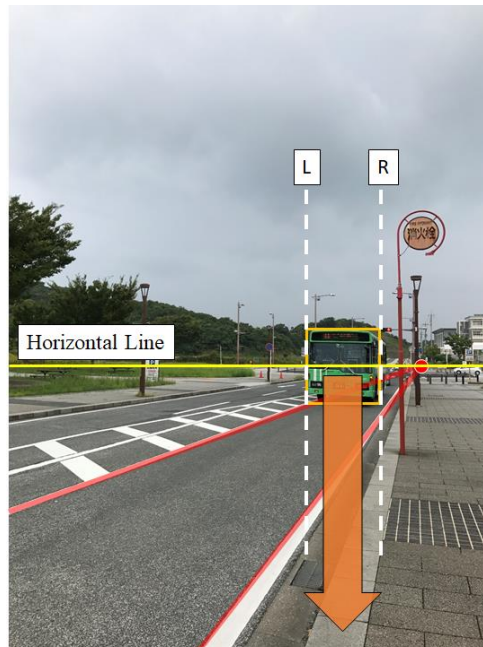


Figure 6.4 Horizon line and oncoming bus.

6.3 Proposed Method for the Obstacle Detection

There are two main processes of obstacle consideration along the road for the proposed system. The first step is to find the obstacle position along the road, and the second step is to estimate the height of the obstacles. For the first step, the combined technique of existing image processing is applied, based on proposed in chapter 4, such as Hough Line Transformation, Uniform Local Binary Pattern, and vertical projection and others. For the second step, the height of obstacles will be estimated, which related to the suitable viewpoint definition in chapter 3. Moreover, the color moment technique will be applied for the second step. Therefore, this section explains whole detail of proposed method for obstacle detection along the road, which the section 6.3.1 will propose the first step of obstacle position detection, and section 6.3.2 will show the method of estimation for height of the obstacles.

6.3.1 The Detection of Obstacle's Position along the Road

The first step of obstacle along the road detection is to find the obstacle position. There are two sub-steps for the proposed method consisting of road area detection and obstacle detection as shown in Figure 6.5.

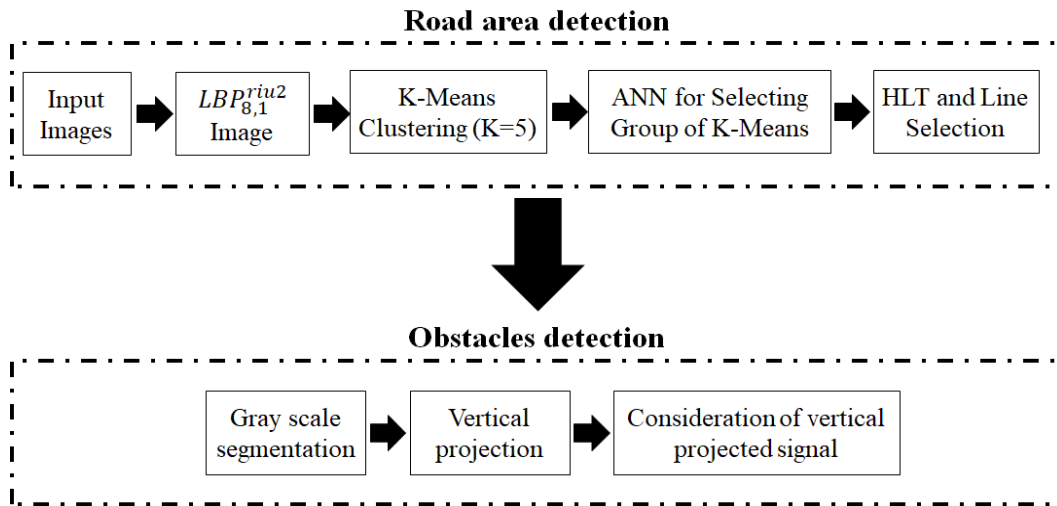


Figure 6.5 Block diagram of the detection of obstacle's position.

6.3.1.1 Road Area Detection

The road area detection is proposed that applied the similar technique as mentioned in chapter 4. However, each step of proposed method will be explained roughly in this section.

1. Input images

All images that are self-collected are acquired in the RGB color format with 800×600 pixel size. Images are taken randomly from the viewpoints on the roadside when users wait for buses.

2. Rotational invariant of uniform local binary patterns ($LBP_{P,R}^{riu2}$)

Road areas often have no certain shape and each scene can be illuminated differently. The $LBP_{P,R}^{riu2}$ is one of the textural features that is robust to rotational texture, as proposed by T. Ojala [41]. This step was applied as same as the proposed method in chapter 4 (section 4.2.1.1).

The $LBP_{8,1}^{riu2}$ technique is applied on a gray scale image as shown in Figure 6.6A. The transformation image of $LBP_{8,1}^{riu2}$ is shown in Figure 6.6B; each 30×30 sub-window size on $LBP_{8,1}^{riu2}$ image is calculated as normalized histogram by ten bins.

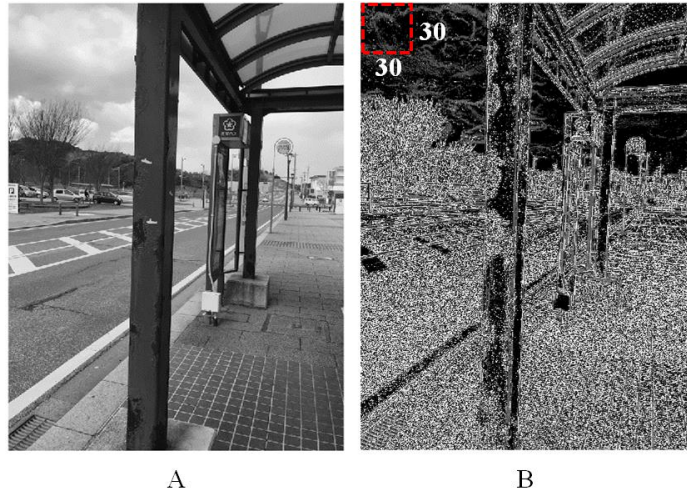


Figure 6.6 The process of image transformation using Rotational invariant of uniform local binary patterns.

A. Gray scale image, B. $LBP_{8,1}^{riu2}$ transformation image.

3. k -means clustering

In order to obtain a similar normalized histogram of each sub-window, k -means clustering technique is used with k set to five and Euclidean distance used as the similarity measure. Figure 6.7 shows outcome of five different labels for each group from the k -means clustering process.



Figure 6.7 Image transformation from k -means clustering.

4. k -means selection using ANN

Figure 6.8 shows the mean of five different labels (k_n) from the k -means process. At the same time, only one label that contains the road area can be selected. Feedforward-backpropagation ANN is used for this selection process. The input layer of the ANN contains ten nodes for ten bins of each histogram. After that, each node of the input layer is multiplied by a different weight value and is summarized in the hidden layer of ANN. Subsequently, the

log-sigmoid function transfers the output value of either 0 or 1 representing whether the label does not contain or does contain the road area respectively.

The ANN is trained with 150 histogram data sets (5 histograms per data set); each data set consists of examples representing both road areas and non-road areas. For the testing step of the ANN, the road area and no road area are represented as 1 and 0 respectively, as shown in Figure 6.8.

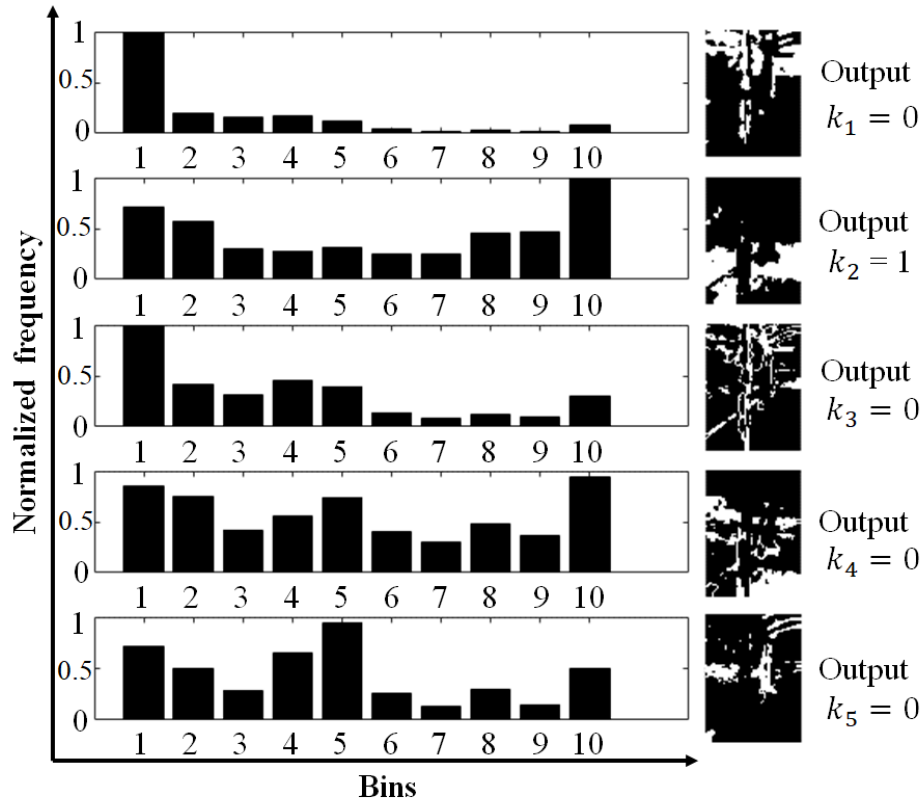


Figure 6.8 Five different histogram representation of $LBP_{8,1}^{riu2}$ using k -means.

5. Hough Line Transformation and line selection

This step is to border the road area using HLT [46-47] for detecting the straight-line markers on the road. As same as the chapter 4 (section 4.2.3.1), the θ values between 20 to 90 degrees was defined for line detection. Figure 6.9A demonstrates the image edge detection that uses Sobel technique, and line detection is applied by using HLT. However, many possible lines and road areas can appear from the previous step.

Although the label that contains the road area is selected, it represents many possible sections of the road area as shown in Figure 6.9B. For this study, two criteria are defined for the initial road area selection. For the first criterion, a group of pixels on the left side and below the image is first selected, as the red area in Figure 6.9B, because the viewpoints for bus-waiting always show up in the road area in this position of the image. The second criterion is applied when the left side of the image does not contain a road area as shown in Figure 6.10A. For example, the selected label in Figure 6.10B shows two groups of pixels and their centroids. In this case, the enclosed group of pixels is selected by the detected lines as shown in Figure 6.10C.

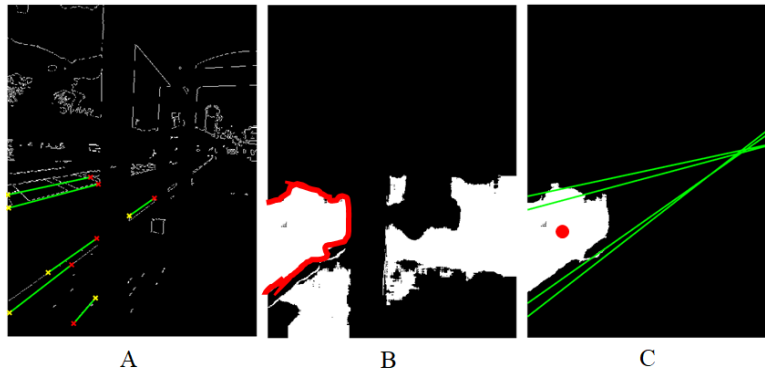


Figure 6.9 Road line detection and initial road area selection (first criterion).

A. Line detection using HLT, B. Selected group of k_n , C. Initial road area and feasible detected lines.

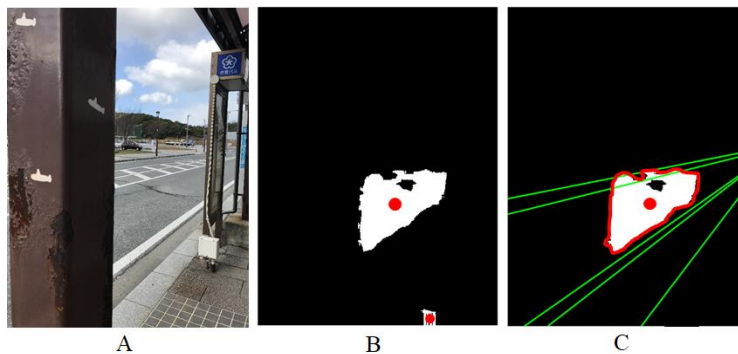


Figure 6.10 Road line detection and initial road area selection (second criterion).

A. Line detection using HLT, B. Selected group of k_n , C. Initial road area and feasible detected lines.

For the line selection, in order to find two actual lines of a road area, this study proposes to use the line touching circle and selects the shortest distance between the centroid of the initial road area and the detected lines. As same as the proposed method in chapter 4 (section 4.2.3.1), the slope lines and r value are calculated as Eq. 4.4 and Eq. 4.5, respectively. The proposed method is applied for each line as shown in Figure 6.11B. Since this step needs two actual lines for the road border, the graph is separated into two areas comprising an upper and a lower centroid; the minimum of r value for each area is selected as shown in Figure 6.11B.

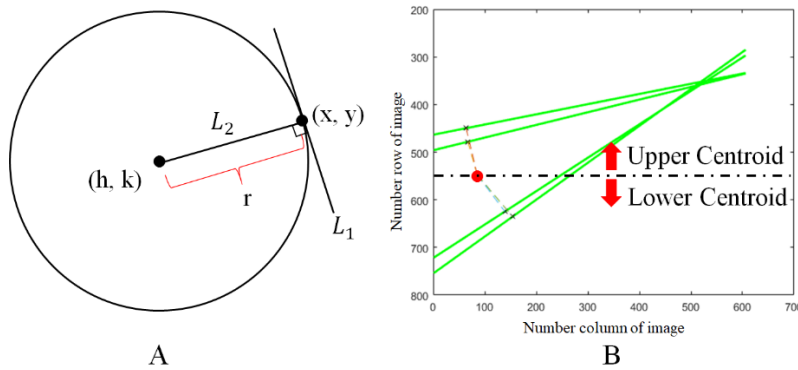


Figure 6.11 The line selection using line touching the circle.

A. Line touching the circle at x, y point, B. Upper and lower centroid separation.

6.3.1.2 Obstacles Detection

There are three steps for finding the position of obstacles along the road, as explained in following;

1. Gray scale segmentation

Up to now, the road area has been defined with two detected lines. Moreover, the initial road area is known from the previous process. This step is region extraction using gray scale segmentation inside the road border. Although gray scale thresholding segmentation is difficult for outdoor scenes, it can be applied for this research. Because the initial road area is known, the gray scale histogram range of the road can be extracted as a reference histogram. Figure 6.12B shows the outcome of segmentation as a binary image where the white pixels are obstacles along the road.

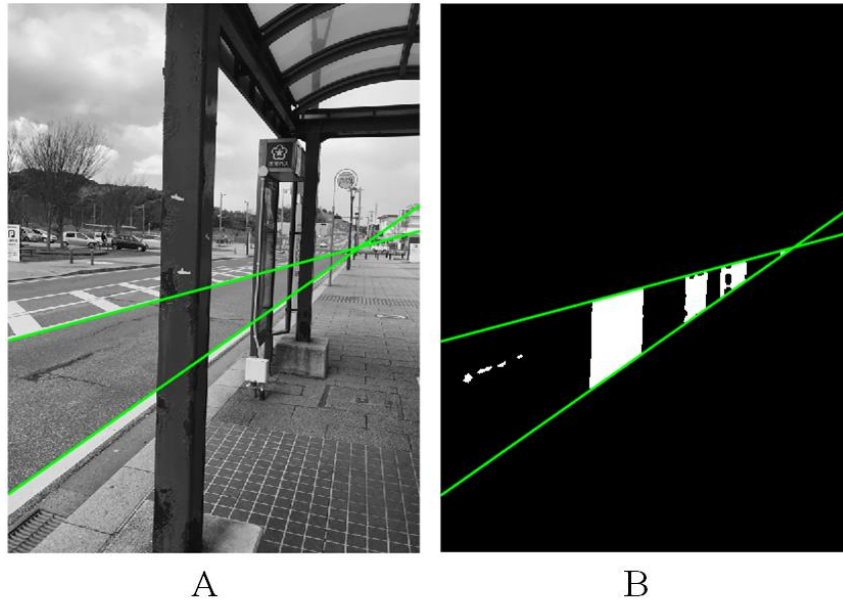


Figure 6.12 Detection of road line and road segmentation.
A. Detection of road lines, B. Binary image of segmentation process.

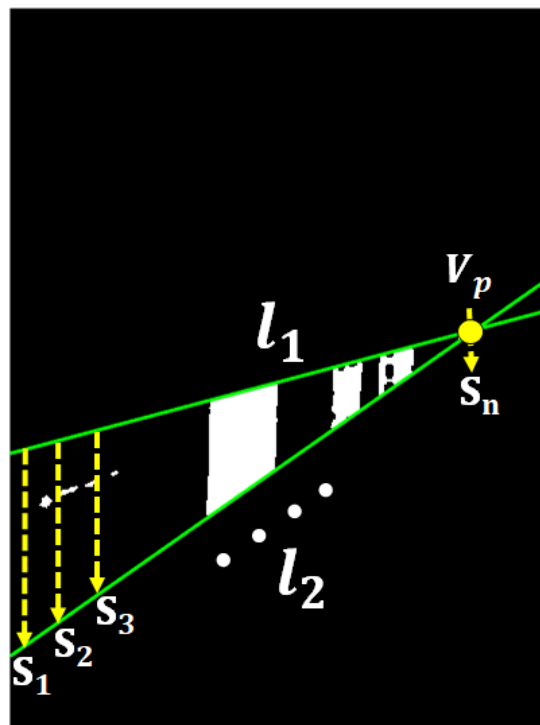


Figure 6.13 Vertical projection of each line scan (S_n) on binary image.

2. Vertical projection

In order to identify the position of obstacles, the summation of vertical projection on binary image is used. In Figure 6.13, each line scan (S_n) summarizes the values in each column from the first column to the vanishing point column (V_p) of the binary image. Since the road border is a line convergence (l_1 and l_2), the summarized value of each line scan is normalized as a percentage of each height (Δl) per line scan. Moreover, the limit of 20% is set for Δl to eliminate noise, because the road border may contain straight-line markers. Therefore, each line scan summation will be set to one or zero, as shown in Eq. 6.1.

$$S_n = \begin{cases} 1; & S_n > 20\% \text{ of } \Delta l \\ 0; & \text{Otherwise} \end{cases} \quad (6.1)$$

3. Consideration of vertical projected signal

The rising and falling edges of the signal can indicate the location of each obstacle and the number of obstacles as shown in Figure 6.14. Figure 6.15 illustrates the example of successful outcomes for obstacle detection along the road using the proposed method.

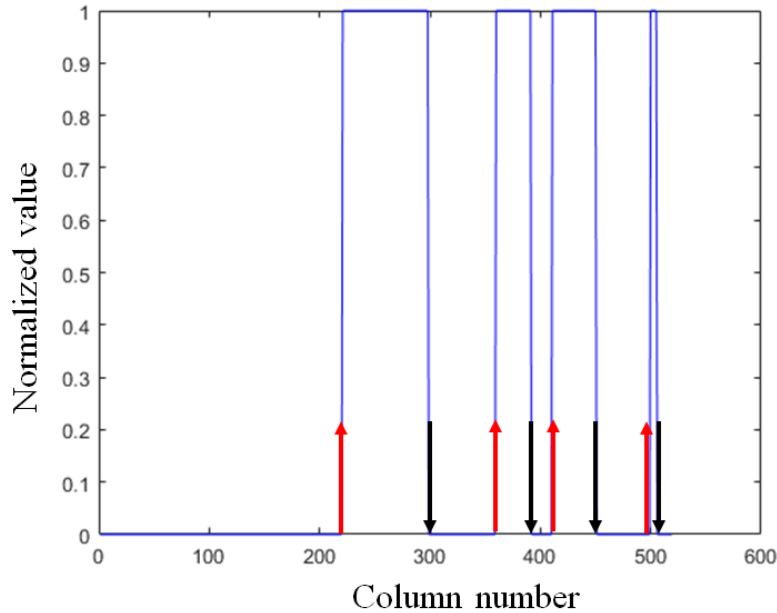


Figure 6.14 Normalized signal of vertical projection.

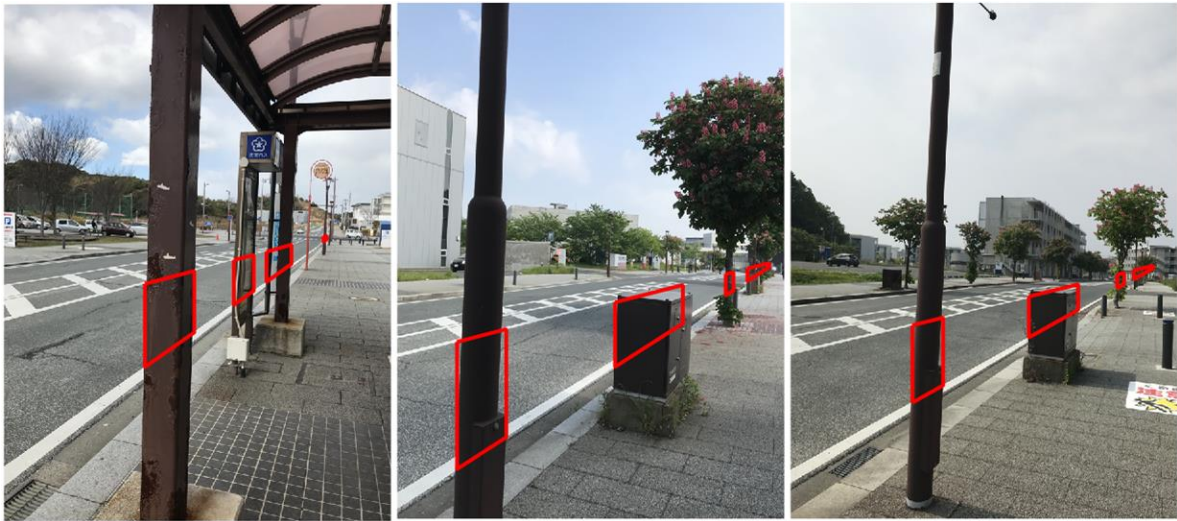


Figure 6.15 Example of obstacle detection using the proposed method.

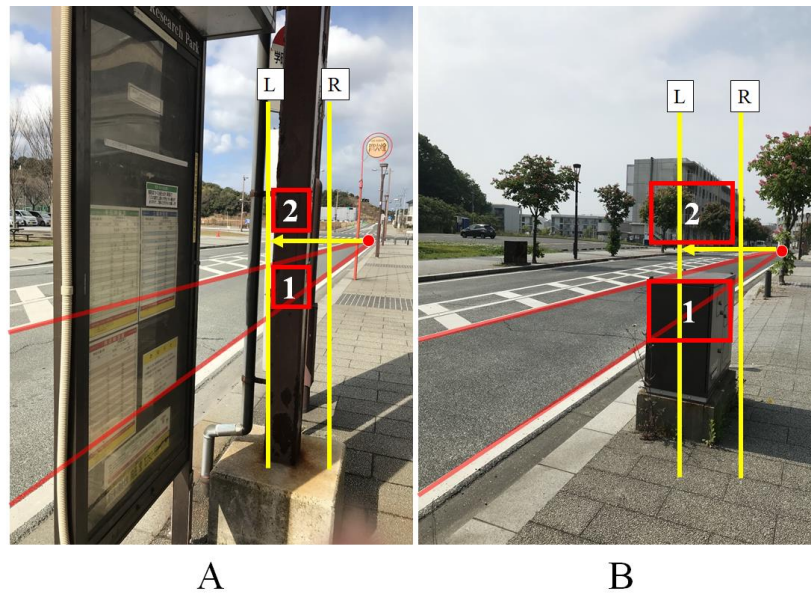


Figure 6.16 Concept for estimating the height of obstacle.

A. Tall obstacle and range of interest, B. Short obstacle and range of interest.

6.3.2 Estimation of Height of Obstacles

The concept for estimating the height of obstacle is to compare areas 1 and 2 outlined with red squares, as shown in Figure 6.16A and Figure 6.16. The system can detect area 1 in the first step of obstacle consideration, as explain in section 6.3.1. In addition, area 2 is an

unknown area that is defined with the same square size as area 1 and located above the horizon line. Our system assumes that the height of the obstacles has to be lower than the horizon line to maintain a suitable viewpoint for bus number detection, as discussed in section 6.2.2. For estimation of height, this study applies the color moment technique [56] that is simply and quickly for calculation.

6.3.2.1 Color Moment Technique

To estimate the height of the obstacles along the road, this study applies a color moment technique. Generally, the color moment technique is used for image retrieval applications, which compares the unknown image and an image in a database using Eq. 6.2. H and I represent the unknown image and the image in the database, respectively. Each i -th color channel is calculated by three moments that consist of mean (E), standard deviation (σ), and skewness (s) as expressed in Eq. 6.3, Eq. 6.4, and Eq. 6.5 with j -th image pixel (p_{ij}). In addition, weight (w) is a parameter that can be adjusted manually depending on the particular application.

$$d(H, I) = \sum_{i=1}^r \{w_{i1}|E_i^1 - E_i^2| + w_{i2}|\sigma_i^1 - \sigma_i^2| + w_{i3}|s_i^1 - s_i^2|\} \quad (6.2)$$

$$E_i = \frac{1}{N} \sum_{j=1}^N p_{ij} \quad (6.3)$$

$$\sigma_i = (\frac{1}{N} \sum_{j=1}^N (p_{ij} - E_i)^2)^{1/2} \quad (6.4)$$

$$s_i = (\frac{1}{N} \sum_{j=1}^N (p_{ij} - E_i)^3)^{1/3} \quad (6.5)$$

Our application must distinguish the pair of images between matching (high obstacle) and non-matching (short obstacle) images, as shown in Figure 6.16A and Figure 6.16B. Thus, this study proposes a simple mathematical model for clustering the category of the pair of images based on the color moment technique.

6.3.2.2 The Concept of Proposed Method for Height Estimation

The concept consists of three main steps, as shown in Figure 6.17. First, the sample square images, which are two categories of matching and non-matching images from the real roadside scenarios, are collected by different sizes of the images depending on the step of obstacle position detection. Then, the color moment in Eq. 6.2 is calculated for each pair of images for both categories. Finally, the standard deviation is calculated for both matching and non-matching data from the color moment outcome. For the ideal graph in Figure 6.17, the calculated color moment of non-matching data will be higher than matching data because there are differences between pairs of images among non-matching data.

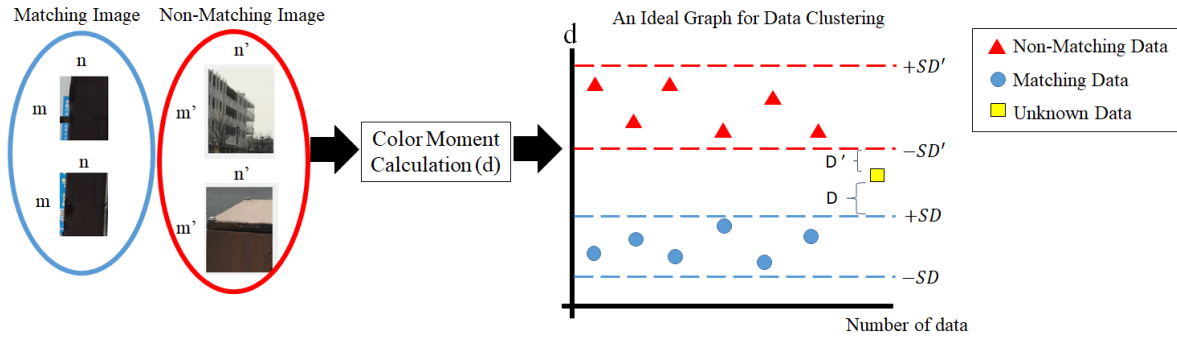


Figure 6.17 The concept of the proposed method and an ideal graph for data clustering.

Moreover, this study proposes a simple method for clustering data between matching and non-matching images using the *mean \pm standard deviation* of each outcome data group. The $\pm SD'$ and $\pm SD$ lines separate the data between two categories, as shown in Figure 6.17, where the $-SD'$ value is higher than the $+SD$ value. Then, when the unknown data are inducted into the system, system can use the criteria with Eq. 6.6 for data assignment. The outcome of the color moment (d) for non-matching data falls on the upper line of $-SD'$, but the d of the matching data is lower than the $+SD$ line. Furthermore, in the case where d falls between the $-SD'$ and $+SD$ lines, as shown in Figure 6.17, it can be calculated by the distance (D, D') between d and $-SD$ and $+SD$ as in Eq. 6.7.

$$\text{case} = \begin{cases} \text{Non matching, } d \geq -SD' | D' < D \\ \text{Matching, } d \leq +SD \end{cases} \quad (6.6)$$

$$D = |d - (-SD')|, D' = |d - (+SD)| \quad (6.7)$$

This section only proposes the feasibility of the concept for clustering the data using the color moment technique. However, it still has some points that we should consider, such as the type of the color model and the weight parameter of the color moment technique, as shown in Eq. 6.2. Therefore, the experiment in section 6.4.2 will focus on this point for optimization.

6.4 Experiments and Results

There are two main experiments corresponding to two steps of obstacle detection. First experiment will show the performance of detection for obstacle's position along the road. Second experiment aims to optimize the performance of obstacle's height estimation, which show the selection of color model and weight selection for color moment technique.

6.4.1 Detection Performance of Obstacle's Position along the Road

This experiment evaluates the efficacy of our proposed static obstacle detection method. Sample images were captured from roadside scenes in non-congested traffic situations and in daytime. In the experiment, 100 images in RGB format and 800×600 image size were tested.

Table 6.1 Confusion Matrix of Obstacle Detection using The Proposed Method.

	Predicted: Yes	Predicted: No	
Actual :Yes	TP = 260	FN = 26	286
Actual :No	FP = 24	TN = 255	279
	284	281	565 (total)

$$\text{Accuracy} = \frac{TP+TN}{\text{Total}} = 0.912 \quad (6.8)$$

$$\text{Recall} = \frac{TP}{\text{Actual Yes}} = 0.909 \quad (6.9)$$

$$\text{Precision} = \frac{TP}{\text{Predicted Yes}} = 0.915 \quad (6.10)$$

$$\text{F - Measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = 0.912 \quad (6.11)$$

Two types of labels were used to distinguish areas with and without obstacles; the total number of instances for both label types was 565. Predicted labels for each instance were compared with their actual labels provided by a manual labeling. Table 6.1 shows the confusion matrix of obstacle detection comprising of the value of true positive (TP), false positive (FP), false negative (FN) and true negative (TN). The experimental results showed the accuracy of 0.912, as shown in Eq. 6.8. The recall and precision values calculated according to Eq. 6.9 and Eq. 6.10 were 0.909 and 0.915 respectively.

6.4.2 Optimization of Color Moment Technique for Obstacle's Height Estimation

To optimize our proposed system, two experiments are considered in this section. First, HSV and RGB color models were selected because they are widely used for image processing. Second, the weight (w) of the color moment in Eq. 6.2 has been optimized in many cases, and each case results in different outcomes.

Each channel of the HSV and RGB color models is calculated by three moments, including the mean (E), standard deviation (σ), and skewness (s) for H and I of each image, which can be represented by matrices in 6.12 and 6.13.

$$H_{hsv} = \begin{bmatrix} h_E & s_E & v_E \\ h_\sigma & s_\sigma & v_\sigma \\ h_s & s_s & v_s \end{bmatrix}, I_{hsv} = \begin{bmatrix} h'_E & s'_E & v'_E \\ h'_\sigma & s'_\sigma & v'_\sigma \\ h'_s & s'_s & v'_s \end{bmatrix} \quad (6.12)$$

$$H_{rgb} = \begin{bmatrix} r_E & g_E & b_E \\ r_\sigma & g_\sigma & b_\sigma \\ r_s & g_s & b_s \end{bmatrix}, I_{rgb} = \begin{bmatrix} r'_E & g'_E & b'_E \\ r'_\sigma & g'_\sigma & b'_\sigma \\ r'_s & g'_s & b'_s \end{bmatrix} \quad (6.13)$$

$$w = \begin{bmatrix} w_1 & w_2 & w_3 \\ w_1 & w_2 & w_3 \\ w_1 & w_2 & w_3 \end{bmatrix} \quad (6.14)$$

For weight assignment, this experiment considers the effectiveness of each channel of the color model using the color moment technique, and thus, the weight (w_1, w_2, w_3) can be assigned as Eq. 6.14. To limit the integer number of weight values 1, 2, and 3 were chosen for our experiment. Further, the weights were rearranged by permutations, excluding the cases of (1, 1, 1), (2, 2, 2), and (3, 3, 3). Therefore, there are 18 cases of weight testing for the optimization process.

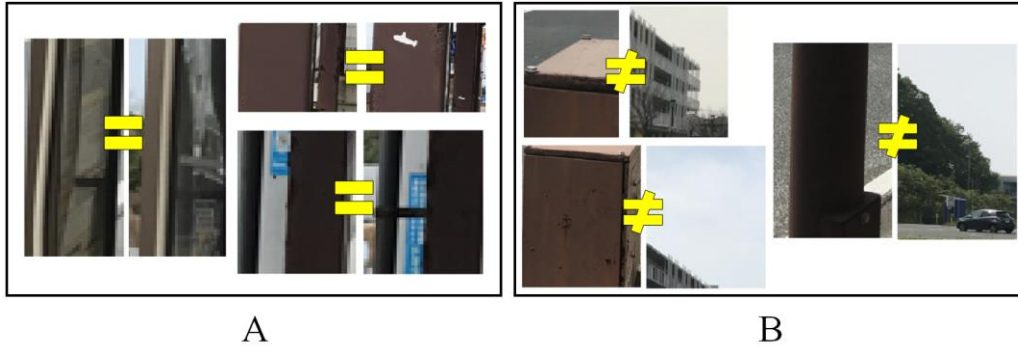


Figure 6.18 Example of matching and non-matching images.

A. Matching images, B. Non-matching images.

In addition, this study provided two different datasets for calculating the criteria and employing the testing process. Each dataset contained 100 RGB and HSV images, where HSV was converted from an RGB color model [57], for matching and non-matching with different square sizes, as shown in Figure 6.18A and Figure 6.18B, respectively, in daylight conditions.

6.4.2.1 HSV and RGB Color Model Selection

According to the color moment technique for image retrieval applications [56], an HSV color model was proposed. However, the RGB color model is also widely used for image processing. Therefore, this experiment compares both models in the experimental analysis.

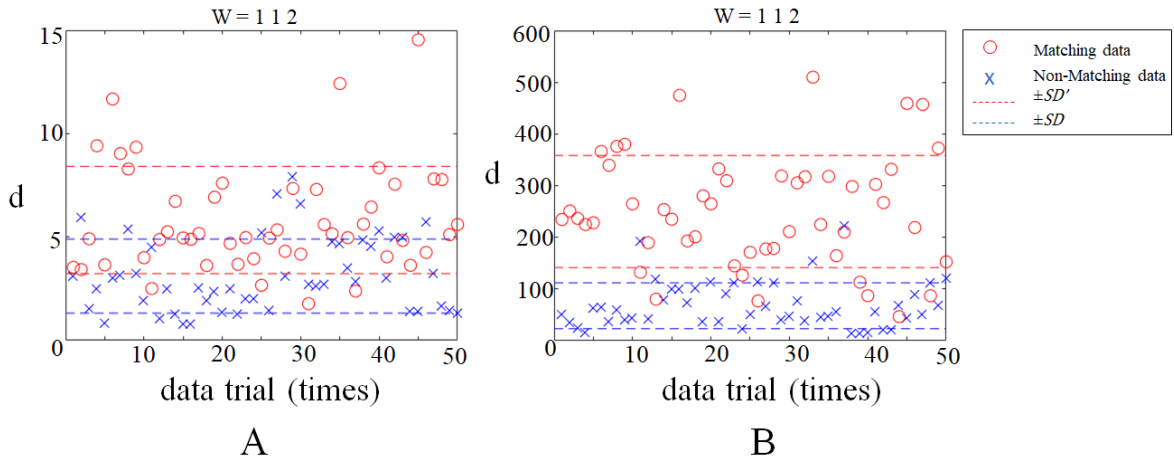


Figure 6.19 Comparison graph for data clustering between HSV and RGB color model using the weight of 1, 1, 2.

A. An example of an HSV color model, B. an example of an RGB color model.

Table 6.2 The $-SD'$ and $+SD$ Comparison between HSV and RGB.

Weights	HSV		RGB	
	$-SD'$	$+SD$	$-SD'$	$+SD$
1, 1, 2	3.23	4.90	107.78	83.57
1, 2, 1	3.18	5.10	131.26	105.20
1, 1, 3	3.94	6.00	111.64	85.22
1, 3, 1	3.77	6.42	154.04	130.10
1, 2, 2	4.08	6.15	131.26	105.20
1, 3, 3	5.63	8.53	161.86	133.39
1, 2, 3	4.86	7.24	139.04	108.47
1, 3, 2	4.76	7.46	157.99	131.73
2, 1, 1	3.01	5.53	172.44	141.03
2, 2, 1	3.93	6.68	203.63	162.36
2, 3, 1	4.67	7.91	232.31	185.04
2, 1, 3	4.70	7.55	180.48	144.18
2, 1, 2	4.79	7.62	176.49	142.60
3, 1, 1	3.47	7.30	240.21	200.63
3, 3, 1	5.36	9.54	303.41	242.76
3, 1, 2	4.45	8.23	244.29	202.19
3, 2, 1	4.50	8.38	272.63	221.18
3, 1, 3	7.15	11.40	248.33	203.75

Given the concept mentioned in section 6.3.2.2, the possible criteria must show $-SD' > +SD$. For example, a comparison of the experimental results of the HSV and RGB color models with weight (1, 1, 2) is presented in Figure 6.19A and Figure 6.19B, respectively. Obviously, the HSV model could not cluster the data between matching and non-matching images, whereas the RGB model performed rather well. Table 6.2 shows the values of $-SD'$ and $+SD$ for the HSV and RGB color models with different weights. Although the weights were changed, the $-SD'$ values were still lower than the $+SD$ values in all cases, although $-SD' > +SD$ was only shown by the RGB color model for all tested cases. Therefore, the RGB color model was selected for optimization in the next step.

6.4.2.2 Weight Selection of Color Moment

Although the RGB color model was chosen as shown in the previous experiment, a set of weights must be selected to optimize performance, and thus, this experiment is referred to as optimization for weight selection.

To optimize the performance of each weight value, a testing dataset of 100 pairs of images was used for matching and non-matching, which is different from previous datasets. Table 6.3 shows the performance of each weight testing, where the highest accuracy average is 85% for weights (1, 1, 2), (1, 1, 3), (1, 3, 3), (1, 2, 3), (2, 3, 1), (2, 1, 2), and (3, 1, 3).

$$H'_{rgb} = \begin{bmatrix} r_E & r_\sigma & r_s \\ g_E & g_\sigma & g_s \\ b_E & b_\sigma & b_s \end{bmatrix}, I'_{rgb} = \begin{bmatrix} r'_E & r'_\sigma & r'_s \\ g'_E & g'_\sigma & g'_s \\ b'_E & b'_\sigma & b'_s \end{bmatrix} \quad (6.15)$$

Although the experimental results displayed a relatively high accuracy, we attempted to further improve the accuracy of our proposed method. The H and I matrices of the RGB model described in Eq. 6.13 were modified by a transpose matrix (H', I'), as shown in Eq. 6.15. However, the same weights for testing were used in the previous experiment. Table 6.4 shows the improved performance by the highest accuracy (86%) in case of (3, 1, 1) and (3, 1, 2) of weight testing.

Table 6.3 Performance of RGB Color Model Testing based on Non-Transpose Matrix (H, I).

Weights	Accuracy (%)		
	Matching images	Non-Matching Images	Average
1, 1, 2	88	82	85
1, 2, 1	86	82	84
1, 1, 3	90	80	85
1, 3, 1	86	80	83
1, 2, 2	86	80	83
1, 3, 3	90	80	85
1, 2, 3	90	80	85
1, 3, 2	86	80	83
2, 1, 1	84	84	84
2, 2, 1	84	84	84
2, 3, 1	86	84	85
2, 1, 3	86	82	84
2, 1, 2	86	82	85
3, 1, 1	84	84	84
3, 3, 1	84	84	84
3, 1, 2	84	84	84
3, 2, 1	84	84	84
3, 1, 3	86	82	85

Table 6.4 Performance of RGB Color Model Testing based on Transpose Matrix (H', I').

Weights	Accuracy (%)		
	Matching images	Non-Matching Images	Average
1, 1, 2	86	82	84
1, 2, 1	90	80	85
1, 1, 3	86	82	84
1, 3, 1	90	78	84
1, 2, 2	90	80	85
1, 3, 3	88	78	83
1, 2, 3	90	80	85
1, 3, 2	88	78	83
2, 1, 1	88	80	84
2, 2, 1	86	82	84
2, 3, 1	88	80	84
2, 1, 3	86	84	85
2, 1, 2	88	82	85
3, 1, 1	88	84	86

Table 6.4 Performance of RGB Color Model Testing based on Transpose Matrix (H' , I'). (cont.)

Weights	Accuracy (%)		
	Matching images	Non-Matching Images	Average
3, 1, 2	88	84	86
3, 2, 1	86	82	84
3, 1, 3	88	82	85

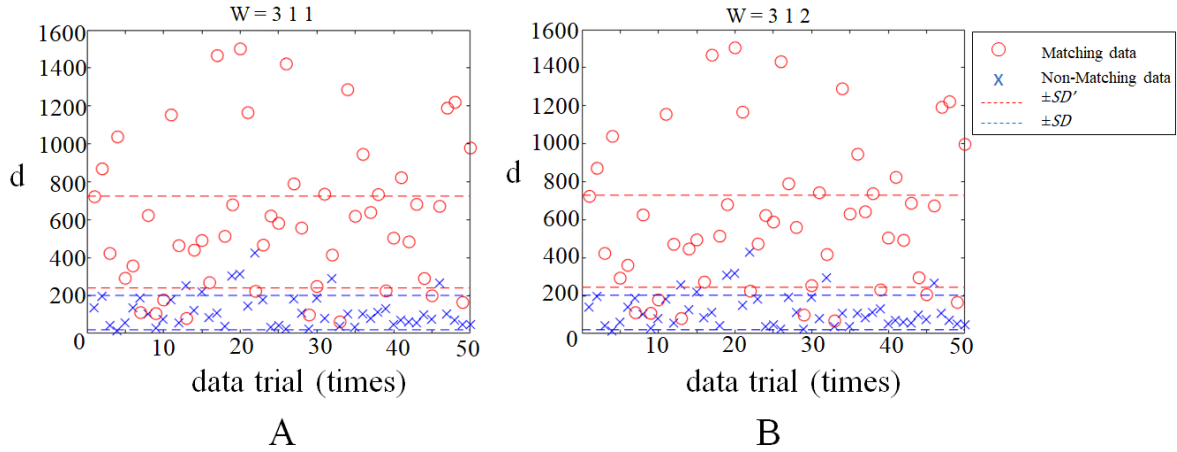


Figure 6.20 Comparison graph of transpose matrix (H' , I') between the weight of 3, 1, 1 and 3, 1, 2.
A. The transpose matrix testing with a weight of 3, 1, 1, B. The transpose matrix testing with a weight of (3, 1, 2).

Herein, weights (3, 1, 1) and (3, 1, 2) were considered particularly because of their highest degree of accuracy. Figure 6.20 shows the plotted testing data and their criteria. In addition, some errors appeared in the plotted graph. For example, some matching data fell below the $\pm SD$ line, and some non-matching data fell above the $-SD'$. To select the optimized weight for our application, the average distance error (\bar{D}_{err}) was measured for each error point (E), as shown in Eq. 6.16. E'_i and E_i represent the error point for matching and non-matching data, respectively.

$$\bar{D}_{err} = \frac{1}{2} \left(\frac{1}{M} (\sum_{i=1}^M |(-SD') - E'_i|) + \frac{1}{N} (\sum_{i=1}^N |(-SD) - E_i|) \right) \quad (6.16)$$

Table 6.5 Comparison of Distance Error Measurements between Weight of (3, 1, 1) and (3, 1, 2).

Weights	Distance error
3, 1, 1	111.12
3, 1, 2	112.32

Table 6.5 shows the outcome of the distance error measurement between weights (3, 1, 1) and (3, 1, 2), where weight (3, 1, 1) displayed an error value lower than weight (3, 1, 2). Therefore, weight (3, 1, 1) of the transpose matrix for the RGB color model was selected for the implementation of the application in a future work.

6.5 Discussion

Since, there were two main experiments as shown in section 6.4, this section will be discussed those experimental results.

6.5.1 Discussion of Obstacle's Position Detection

Based on the experimental result of obstacle's position detection, the performance of the proposed method was 0.912 for the F-measure value representing the accuracy rate. However, some error cases appeared when obstacle size was very small at the far position in a perspective view, thus we think it should be eliminated as the noise area. The error rate was 0.086, as shown in Eq. 6.17.

$$\text{Error Rate} = \frac{FP+FN}{Total} = 0.086 \quad (6.17)$$

The proposed method showed a high performance for the step of static obstacle detection along the road. However, this combined method was applied to the gray scale image thresholding for outdoor scenes in only daytime condition.

6.5.2 Discussion of Optimization Process for Obstacle's Height Estimation

According to the experimental result of obstacle's height estimation step, two experiments consisting of color model selection and the weight of color moment selection

were performed. In addition, this study also modified matrices H and I by the transpose matrices H' and I' to improve the accuracy of the proposed method.

On the basis of the experimental results, the RGB color model was able to distinguish ($-SD' > +SD$) between the data group of the matching and non-matching images in all cases of weight testing, but the results of HSV color model testing were totally dissimilar compared to the RGB modeling results.

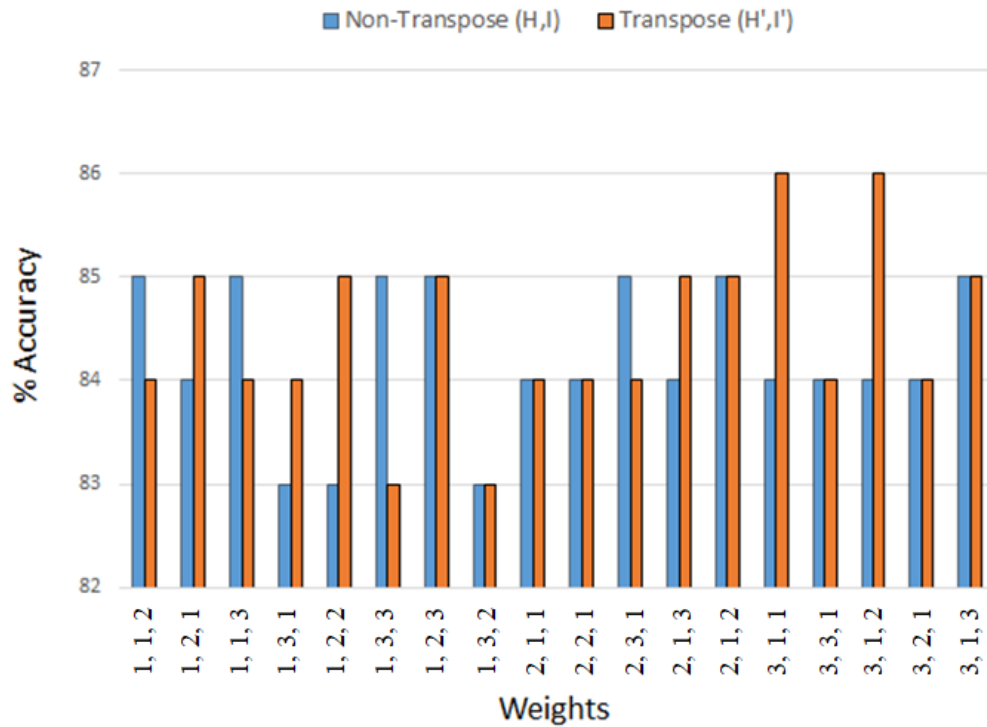


Figure 6.21 Accuracy comparison between non-transpose (H, I) and transpose (H', I') matrices.

Furthermore, 18 cases of weight were tested using the RGB color model. First, matrices H and I in Eq. 6.13 were calculated, and the results showed the highest accuracy of 85% for many cases of weight testing. Second, the transpose matrices H' and I' in Eq. 6.15 were used to further improve the accuracy, and the highest accuracy was 1% higher than the results of the H and I matrix testing. Figure 6.21 compares the accuracies of the non-transpose and transpose matrices, which resulted in the improvement in many cases, as shown for weights (1, 2, 1), (1, 3, 1), (1, 2, 2), (2, 1, 3), (3, 1, 1), and (3, 1, 2). Particularly in weight

cases of (3, 1, 1) and (3, 1, 2), which showed the highest accuracy of 86%, use of the transpose matrices resulted in performance that was better than all cases employing non-transpose matrices. As a result, the mean (μ) moment for each channel of RGB had a significant effect on performance. Although the standard deviation (σ) and skewness (s) were less significant than the mean, the distance error increased when the skewness was defined by 2, as shown in Table 6.5.

Figure 6.22 shows the detection of obstacles along the road by our proposed method in real-world daytime conditions. The red boxes, shown in Figure 6.22A and Figure 6.22B, identify the obstacle height and deem the location as unsuitable to wait for the bus, and the green box in Figure 6.22C identifies a short obstacle along the road that does not obscure the suitable viewpoint and location to wait for the bus. However, this proposed method is still limited with respect to its application in nightlight conditions because of extreme darkness and differences in illumination.



Figure 6.22 Example of obstacles detection using our proposed method.

A. and B. The detection of height obstacles with red boxes, C. The detection of a short obstacle with a green box.

6.6 Summary

This chapter presents the obstacle detection along the road, which will help the blind people to know the suitable viewpoint of the bus-waiting. Especially, the situation of non-congested traffic is considered.

Since, some obstacles along the road may obscure the suitable viewpoint while waiting for the bus, thus the important properties of obstacle are explained in this chapter. There are two properties of unsuitable obstacles along the road related to the suitable viewpoint definition. First property is the position of obstacle in the image that can obscure the bus number of oncoming bus. Second property is proposed as the height of obstacle, which related to the horizon line of perspective view.

Moreover, the proposed method of obstacle detection consists of two main steps. First, the obstacle position will be detected, which has two sub-steps of the road area detection and obstacle position detection. Second, the height of obstacles will be estimated, using the color moment technique.

For the experiments, two main steps of obstacle detection are tested, especially the images are collected in daytime condition. First experiment shows the performance of obstacle position detection in term of F-measure by 0.912. Second experiment shows the selection of color model, and optimization of weight for color moment technique. The experimental results can conclude that the RGB color model provided feasible results, whereby $-SD' > +SD$ was shown by all weights tested (18 cases). Moreover, all cases of weight testing for the RGB model were optimized by accuracy measurements. In addition, this study compares the non-transpose and transpose matrices of H and I for the proposed color moment technique, and the results indicated that the transpose matrix performed the best with an accuracy of 86%.

Finally, the RGB color model with a weight of (3, 1, 1) for the transpose matrix was selected for application in our future research. However, it still has the limitation of not being usable in nighttime conditions, which we are planning to consider in our future work.

This page intentionally left blank

7 Conclusions and Future Work

7.1 Conclusion

The bus identification system is still important for assisting blind people while waiting for the bus. Especially the underdeveloped and developing countries where have the poor system of the public bus service. This study proposed a part of bus identification, which is the consideration of viewpoints before the bus detection. The suitable viewpoint of bus-waiting is a crucial step of the system to let blind know the possibility for detecting the oncoming bus.

Since, there was no any previous research proposing the viewpoint classification while waiting for the bus. The definition of suitable viewpoint of bus-waiting for helping blind people was proposed in chapter 3. Then, the consideration of viewpoints while waiting for the bus was presented, which there were three main studies as follows:

- 1) The situation of non-congested traffic was considered in chapter 4. This study proposed the classification of viewpoint using the combined method of existing image processing techniques. To segment and provide the road area features, the Rotational Invariant of Uniform Local Binary Pattern ($LBP_{P,R}^{riu2}$) was main used. Moreover, other technique such as k -means, Hough Line Transformation, and line touching circle were applied for this chapter. Then, the Artificial Neural Network was used for classification step. For the experiment, 8 conditions of weather under daytime and nighttime were tested as shown in Table 4.5, For the final result of this study, 98.56% of accuracy for suitable viewpoint classification.
- 2) The situation of congested traffic was also considered in chapter 5. Since, the road feature could not perform for this situation because many cars obscured the road area. This study proposed the car distribution in the image for viewpoint classification. First, YOLOv2 technique was applied to detect the cars in image with the parameter setting in Table 5.1. Then, the 19 feasible features of car distribution were extracted. For the experiment, the performance of classification, both of daytime and nighttime, was optimized by testing various classifiers and different feature selection. According

to the results in Table 5.9, a random forest classifier and 17 features provided the highest accuracy by 86%.

- 3) Although, the situation of non- and congested traffic were considered, the real situation might appear obstacles that can obscure the suitable viewpoint. The obstacle detection was proposed in chapter 6, especially the case of non-congested traffic in daytime. There were two main properties of unsuitable obstacles along the road related to definition of suitable viewpoint of the bus-waiting. First property is the obstacle's position in image. Second property is obstacle's height. Therefore, this study proposed the two main steps of obstacle detection corresponding to the unsuitable obstacle properties. First step, the position of obstacle was detected using the combined method and vertical projection technique. Then, the color moment technique was applied for estimating the height of obstacle. According to experimental result, 0.918 of F-measure (91.2% for accuracy) showed the performance of obstacle's position detection, and the performance of height estimation was optimized and showed the accuracy by 86%.

However, there is a concern about dangerous for blind users, who take the image very close to the road area. They might get some accidents from vehicles on the road. Therefore, they have to be careful themselves for using this system.

7.2 Future Work

In order to improve and realize the classification system of viewpoint while waiting for the bus, some points of system should be considered in future work as shown in Figure 7.1.

Although, this dissertation proposed the classification of viewpoints under clearly situation between non-congested traffic and congested traffic, the algorithm for selecting the situations automatically do not propose. To distinguish these two situations, we may consider and set the criterion of ratio between the percentages of road area and number of cars in the image.

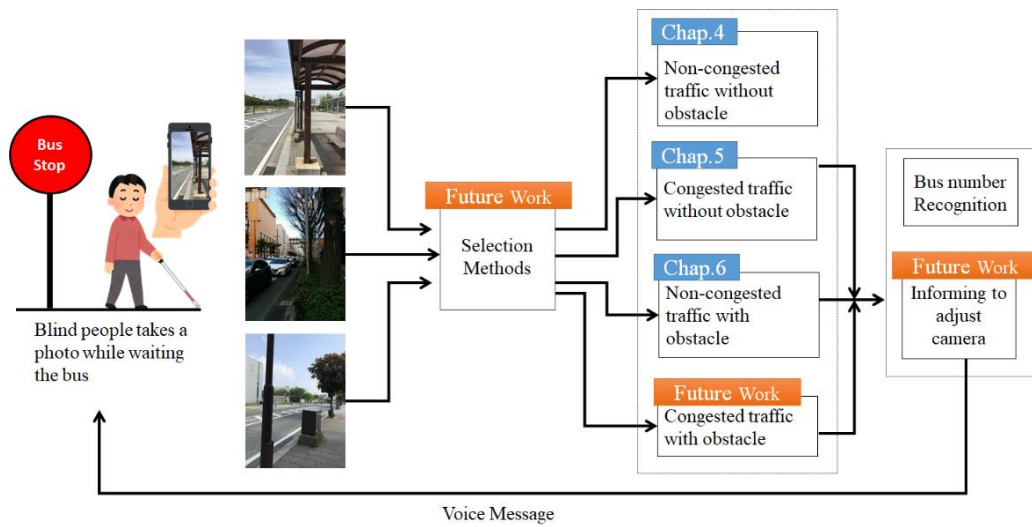


Figure 7.1 Overview of proposed system and future works

Moreover, the detection of obstacles along the road, especially under nighttime, is still extremely difficult to use the same technique as proposed in chapter 6. For future work, we aim to solve this situation with different methods. For example, adjusting the weight of color moment technique for nighttime or using the more complex algorithms.

In addition, a case of obstacles detection in situation of congested traffic do not solve for this research. Since, the problem of vanishing point on perspective view is still difficult for estimation. However, this situation is necessary in order to realize the system of viewpoint classification in the future.

A case of curved road did not consider in this research because the definition of this research defined the viewpoints based on straight road, as a perspective view, as shown in Figure 7.2A. However, a case of much curved road at some bus stop areas as shown in Figure 7.2B, the classification of viewpoint will be confused. Therefore, the curved road viewpoint will be solved in future.

To let the blind users adjust their smartphone position when the viewpoint is unsuitable, the notification solution is very crucial. Since, the blind user cannot see anything; the sound message can guide them to adjust their smartphone position. However, the comfortable and suitable system of sound guidance is an important future experiment to realize the complete system for blind users.

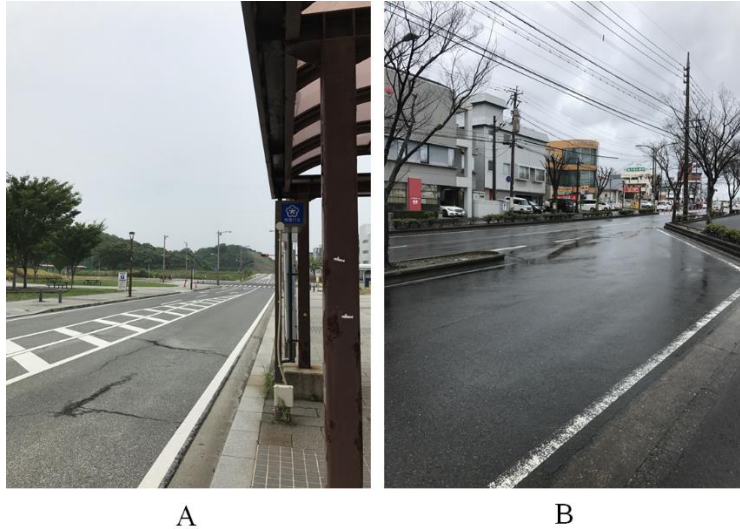


Figure 7.2 Example image of straight and curved roads.
A. straight road, B. and C. curved road

To realize the system in the real-time of this viewpoint classification, we plan to implement for movie image on smartphone after completing all future works of system as above- mentioned. Based on the proposed methods and experimental results, we are quite sure that the system will perform well for the real situation to help blind people.

References

- 1 World Health Organization. [Online], <http://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>, accessed on 11 October 2018.
- 2 Wikipedia, “Visual Impairment” [Online], https://en.wikipedia.org/wiki/Visual_impairment, accessed on 3 May 2019.
- 3 WeWALK, “Smart cane helps visually impaired see obstacles – CNN [Online], <https://wewalk.io/>, accessed on 1 January 2018.
- 4 OrCam Myeye2 for blind and visually impaired [Online], <https://www.orcam.com/en/myeye2/>, accessed on 10 May 2019.
- 5 Maureen Duffy, “Google Glass Applications for Blind and Visually Impaired Users” [Online], <https://www.visionaware.org/blog/visionaware-blog/google-glass-applications-for-blind-and-visually-impaired-users/12>, accessed on 8 May 2013.
- 6 Team DiyKerala, “SIGHT: For the Blind” [Online], <https://www.hackster.io/diykerala/sight-for-the-blind-c1e1b9>, accessed on 30 October 2017.
- 7 iMore, “How to use VoiceOver on iPhone and Ipad” [Online], <https://www.imore.com/how-use-voiceover-iphone-and-ipad>, accessed on 3 February 2019.
- 8 Santos, E. A. B., “Design of an interactive system for city bus transport and visually impaired people using wireless communication, smartphone and embedded system”, In *Microwave and Optoelectronics Conference (IMOC), SBMO/IEEE MTT-S International*, pp. 1-5, November, 2015.
- 9 El Alamy, L., Lhaddad, S., Maalal, S., Taybi, Y., & Salih-Alj, Y, “Bus identification system for visually impaired person”, In *6th International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST)*, pp. 13-17, September, 2012.
- 10 Venard, O., Baudoin, G., & Uzan, G., “Field experimentation of the RAMPE interactive auditive information system for the mobility of blind people in public

- transport: Final evaluation”, In *9th IEEE International Conference on Intelligent Transport Systems Telecommunications (ITST)*, pp. 558-563, October, 2009.
- 11 Noor, M. Z. H., Ismail, I., & Saaïd, M. F. “Bus detection device for the blind using RFID application” In *5th IEEE International Colloquium on Signal Processing & Its Applications (CSPA)*, pp. 247-249, March, 2009.
 - 12 Al Kalbani, J., Suwailam, R. B., Al Yafai, A., Al Abri, D., & Awadalla, M., “Bus detection system for blind people using RFID”, In *8th IEEE GCC Conference and Exhibition (GCCCE)*, pp. 1-6, February, 2015.
 - 13 Zhou, P., Zheng, Y., & Li, M., “How long to wait?: predicting bus arrival time with mobile phone based participatory sensing” In *Proceedings of the 10th international conference on Mobile systems, applications, and services*, pp. 379-392, June, 2012.
 - 14 *Digital Age Magazine* [Thai], “Introduction of ViaBus Application”, ISSUE 216, December, 2017.
 - 15 Wongta, P., Kobchaisawat, T., & Chalidabhongse, T. H., “An automatic bus route number recognition”, In *13th IEEE International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pp. 1-6, July, 2016.
 - 16 Cheng, C. C., Tsai, C. M., & Yeh, Z. M., “Detection of bus route number via motion and YCbCr features”, In *IEEE International Symposium on Computer, Consumer and Control (IS3C)*, pp. 31-34, June, 2014.
 - 17 Lee, D., Yoon, H., Park, C., Kim, J., & Park, C. H., “Automatic number recognition for bus route information aid for the visually-impaired”, In *10th IEEE International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pp. 280-284, October, 2013.
 - 18 Guida, C., Comanducci, D., & Colombo, C., “Automatic bus line number localization and recognition on mobile phones—a computer vision aid for the visually impaired”, In *International Conference on Image Analysis and Processing* Springer, Berlin, Heidelberg, pp. 323-332, September, 2011.
 - 19 Pan, H., Yi, C., & Tian, Y., “A primary travelling assistant system of bus detection and recognition for visually impaired people”, In *IEEE International Conference*

- on *Multimedia and Expo Workshops (ICMEW)*, pp. 1-6, July, 2013.
- 20 American Foundation for the Blind (AFB). Accessible Mass Transit. [Online], <http://www.afb.org/info/living-with-vision-loss/getting-around/accessible-mass-transit/235>, accessed on 3 December 2018.
 - 21 Hersh, M., & Johnson, M. A. (Eds.) “*Assistive technology for visually impaired and blind people*”, In Springer Science & Business Media, 2010.
 - 22 Markiewicz, M., & Skomorowski, M, “Public transport information system for visually impaired and blind people”, In *International Conference on Transport Systems Telematics* Springer, Berlin, Heidelberg, pp. 271-277, October, 2010.
 - 23 Poliński, J., & Ochociński, K., “Tactile Graphics at Railway Stations—an Important Source of Information for Blind and Visually Impaired Travellers”, In *Problemy Kolejnictwa*, 2017.
 - 24 CIVITAS. “Case study public transport information for blind and partially sighted people in Brighton and Hove”, In *ICLEI-Local Governments for Sustainability*, Freiburg, Germany, 2015.
 - 25 Sánchez, J., & Oyarzún, C., “Mobile audio assistance in bus transportation for the blind”, In *International Journal on Disability and Human Development*, Vol.10, No. 4, pp. 365-371, 2011.
 - 26 Watcharin, T., & Chikamune, W., “Classification of Viewpoints Related to Bus-Waiting for the Assistance of Blind People”, In *International Journal of New Technology and Research (IJNTR)*, Vol. 4, No. 9, pp. 43-52, 2018.
 - 27 Watcharin, T., Masashi, N., Kodai, K., & Chikamune, W., “Viewpoint Classification for the Bus-waiting Blinds in Congested Traffic Environment”, In *International Journal of Fuzzy Logic and Intelligent Systems*, Vol. 19, No. 1, , pp. 48-58, March, 2019.
 - 28 Watcharin, T., & Chikamune, W., “Obstacle Height Estimation Related to Suitable Viewpoint while Waiting for the Bus using Color Moment Technique”, In *International Journal of Machine Learning and Computing*, 2019. [in press]
 - 29 Howett, D., “Television Innovations: 50 Technological Developments: a Personal

- Selection”, In *Kelly Publications*, 2006.
- 30 Wikipedia, “The free encyclopedia. Panning (camera)”, [online], [https://en.wikipedia.org/wiki/Panning_\(camera\)](https://en.wikipedia.org/wiki/Panning_(camera)), Access on: 30 November 2018.
 - 31 Dailynews [Thai]. [online news], <https://www.dailynews.co.th/economic/663713>, accessed on 30 November 2018.
 - 32 Department of Land Transport [Thai]. [online]. <https://www.dlt.go.th/th/>, 2018.
 - 33 Ball and Nok blog [Thai]. [online], <http://ball-nok.blogspot.com/2011/05/2553.html>, accessed on 18 May 2010.
 - 34 Bangkok Mass Transit Authority (BMTA) [Thai]. [online]. <http://www.bmta.co.th/th/bus-lines>, accessed on 3 January 2018.
 - 35 Wikipedia, “Bangkok Mass Transit Authority, BMTA” [Thai], December 2018.
 - 36 Ebay. “VICI LX1332B Accuracy Digital Lux meter Light Meter 1~200,000 Lux Luminometer”, <https://www.ebay.com/itm/VICI-LX1332B-Accuracy-Digital-Luxmeter-Light-Meter-1-200-000-Lux-Luminometer-/292481725081>.
 - 37 Smith, R., “An overview of the Tesseract OCR engine”, In *IEEE International Conference on Document Analysis and Recognition*, Vol. 2, pp. 629-633, September, 2007.
 - 38 Smith, R., Antonova, D., & Lee, D. S., “Adapting the Tesseract open source OCR engine for multilingual OCR”, In *Proceedings of the International Workshop on Multilingual OCR*, pp. 1, July, 2009.
 - 39 Smith, R. W., “Hybrid page layout analysis via tab-stop detection”, In *10th IEEE International Conference on Document Analysis and Recognition*. pp. 241-245, July, 2009.
 - 40 Modi, H., & Parikh, M. C., “A review on optical character recognition techniques”, In *Int J Comput Appl*, Vol. 160, No. 6, pp. 20-24, 2017.
 - 41 Ojala, T., Pietikainen, M., & Maenpaa, T., “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns”, In *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 24, No. 7, pp. 971-987, 2002.
 - 42 Arthur, D., & Vassilvitskii, S., “*k*-means++: The advantages of careful seeding”,

- In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, Society for Industrial and Applied Mathematics, pp. 1027-1035, January, 2007.
- 43 Luke, D., “What is an artificial neural network? Here’s everything you need to know” [online], <https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network>, accessed on 1 May 2019.
 - 44 Rojas, R., “Neural Network. Springer-Verlag”, Chapter 7: The Backpropagation Algorithm”, 1996.
 - 45 Hough, P. V., *U.S. Patent No. 3,069,654*. Washington, DC: U.S. Patent and Trademark Office, 1962.
 - 46 Duda, R. O., & Hart, P. E., “Use of the Hough transformation to detect lines and curves in pictures”, In *SRI INTERNATIONAL MENLO PARK CA ARTIFICIAL INTELLIGENCE CENTER*, No. SRI-TN-36, 1971.
 - 47 Ballard, D. H., “Generalizing the Hough transform to detect arbitrary shapes” In *Pattern recognition*, Vol. 13, No. 2, pp. 111-122, 1981.
 - 48 Fisher, R., Perkins, S., Walker A. & Wolfart E. Sobel Edge Detection [Online], <https://homepages.inf.ed.ac.uk/rbf/HIPR2/sobel.htm>, 2003.
 - 49 Adams, R., & Bischof, L., “Seeded region growing”, In *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 16, No. 6, pp. 641-647, 1994.
 - 50 Fan, J., Zeng, G., Body, M., & Hacid, M. S., “Seeded region growing: an extensive and comparative study”, In *Pattern recognition letters*, Vol. 26, No. 8, pp. 1139-1156, 2005.
 - 51 Redmon, J., Divvala, S., Girshick, R., & Farhadi, A., “You only look once: Unified, real-time object detection”, In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779-788, 2016.
 - 52 Redmon, J., & Farhadi, A., “YOLO9000: better, faster”, In stronger. *arXiv preprint*, 2017.
 - 53 Noda, M., Miyamoto, H., Tangsuksant, W., Kitagawa, K., & Wada, C., “Basic Study on Viewpoints Classification Method using Car Distribution on the Road”, In *IEEE*

International Conference on Information and Communication Technology Robotics (ICT-ROBOT), pp. 1-3. September, 2018.

- 54 WEKA: Weka3.8. “Data mining software in Java”. Available: <https://www.cs.waikato.ac.nz/ml/weka>
- 55 Cripky-ghee, “Crispy’s Perspective Basics and Tips-Part 1: Horizon Line” [Online] Available: <http://crispy-ghee.tumblr.com/post/37530860486/crispys-perspective-basics-and-tips-part-1>, accessed on 8 December 2012.
- 56 Stricker, M. A., & Orengo, M., “Similarity of color images. In *Storage and Retrieval for Image and Video Databases III and International Society for Optics and Photonics*”, Vol. 2420, pp. 381-393, March, 1995.
- 57 Smith, A. R., “Color gamut transform pairs”, In *ACM Siggraph Computer Graphics*”, Vol.12, No. 3, pp. 12-19, 1978.

Appendix Published Papers

Journals

- 1 Watcharin Tangsuksant, Chikamune Wada, “*Obstacle Height Estimation Related to Suitable Viewpoint while Waiting for the Bus using Color Moment Technique*”, In International Journal of Machine Learning and Computing. **[in press]**
- 2 Watcharin Tangsuksant, Masashi Noda, Kodai Kitagawa, Chikamune Wada, “*Viewpoint Classification for the Bus-waiting Blinds in Congested Traffic Environment*”, In International Journal of Fuzzy Logic and Intelligent Systems, Volume 19, Issue 1, Mar., 2019, pp. 48-58.
- 3 Watcharin Tangsuksant, Chikamune Wada, “*Classification of Viewpoints Related to Bus-Waiting for the Assistance of Blind People*”, In International Journal of New Technology and Research, Volume 4, Issue 9, Sept. 2018, pp. 43-52.

Conferences

- 1 Watcharin Tangsuksant, Chikamune Wada, “*Comparison of HSV and RGB Color Models using Color Moment for Obstacles Matching along the Road*”, In the 6th International Symposium on Applied Engineering and Sciences, Kitakyushu, Japan, 15-16 Dec, 2018.
- 2 Watcharin Tangsuksant, Chikamune Wada, “*Obstacle Detection along the Road with a Combined Method*”, In the 11th Biomedical Medical International Conference, ChiangMai, Thailand, 21-24 Nov, 2018.
- 3 Masashi Noda, Hiroyuki Miyamoto, Watcharin Tangsuksant, Kodai Kitagawa, Chikamune Wada, “*Basic Study on Viewpoints Classification Method using Car Distribution on the Road*”, In the International Conference on Information and Communication Technology Robotics, Busan, Korea (South), 6-8 Sept, 2018.
- 4 Watcharin Tangsuksant, Chikamune Wada, “*New Combination Method for Road Features Extraction*”, In the International Conference on Information and Communication Technology Robotics, Kagoshima, Japan, 25-26 Nov, 2017.